Effects of Race-Neutral Affirmative Action: Evidence from Academically Selective High Schools *

Rene A. Crespin† Tyler E. Powell‡

February 2026

Abstract

Affirmative action (AA) remains one of the most divisive policy debates in the United States. We provide the first national empirical evidence on how race-neutral AA admissions policies affect the demographic and academic composition of academically selective high schools (ASHSs). To do this, we assemble a novel dataset linking ASHS admissions policies from 2010 onward to national school-cohort records and exploit the staggered implementation of race-neutral AA admissions using event-study and difference-in-differences designs. These reforms increase low-income and underrepresented minority (URM) enrollment, driven largely by declines in White and "other race" shares, with little change in Asian enrollment. We find that average test score declines are modest, with treated cohorts exhibiting roughly a 2% decline in average ACT/SAT performance. To shed light on this effect, we examine two potential mechanisms: declines in incoming academic preparation and increases in race-based harassment/bullying. Finally, we benchmark our results relative to counterfactual/simulation-based predictions that dominate prior work and discuss the potential academic consequences of eliminating SES and URM representation gaps at ASHSs across the country.

*We thank Stacy Dickert-Conlin, Todd Elder, Maria Fitzpatrick, Steven Haider, Scott Imberman, Michael Lovenheim, and Evan Riehl for helpful comments and discussions. Madeleine Smith provided excellent research assistance. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B200009 to Michigan State University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

†Department of Economics, Michigan State University; crespin@msu.edu.

‡Department of Economics, Michigan State University; powel282@msu.edu

Introduction

Affirmative action (AA) in elite school admissions remains one of America's most divisive policy debates. Selective educational institutions offer students access to competitive, well-resourced learning environments that are populated with high-achieving peers (Chetty et al., 2023; Dobbie & Fryer, 2014; Finn & Hockett, 2012). Persistent disparities in access to these institutions raise important questions about how to balance rigorous selectivity with socioeconomic and racial diversity, all while minimizing academic mismatch (Arcidiacono & Lovenheim, 2016). Although recent national discourse has focused on college admissions, K–12 districts with academically selective public schools have had similarly contentious debates over the past several decades.[1]

Selective K–12 schools are often regarded as a source of pride and a means to attract and maintain engagement with middle- and upper-class families in public school districts, offering advanced courses and extracurriculars typically not available in traditional schools (Finn & Hockett, 2012; Hoxby, 2003). Elite colleges often view academically selective high schools (ASHSs) as reliable "feeders" due to their rigorous coursework (Hoxby & Avery, 2013). Moreover, these schools frequently play a role in families' educational decisions, with over 60% of eighth graders applying for admission in some districts (Feerick Center for Social Justice, 2019; Perez, 2018). Consequently, administrators are incentivized to uphold their schools' strong academic reputations while simultaneously addressing concerns about their racial and socioeconomic diversity (Cestau et al., 2017; Ellison & Pathak, 2021). In response to criticism of the lack of diversity at selective schools and limitations on the use of race for admissions, an increasing number of districts have begun to use race-neutral AA policies over the past decade. Despite the growing interest in reforming admissions, research on the potential impacts on school diversity and academic performance has been limited to simulation-based studies without real world policy variation.

In this paper, we provide the first empirical evidence on the effects of race-neutral AA admissions policies on the socioeconomic/racial diversity and academic performance at ASHSs across the United States, relative to status quo ASHS admissions. While most selective public schools do not change their admissions policies frequently, if at all, we leverage geographic and temporal variation in observed admissions policy changes to analyze impacts.

We create a novel national dataset of selective public school admissions policies and their changes over time. We first gather data on the availability of ASHSs since 2010. Next,

[1] In this paper, we use the general term academically selective public schools to refer to any public school that uses competitive admissions. In practice, this type of public school takes various names across districts. In Boston, public schools with academically-competitive admissions are Exam Schools; in Chicago, they are Selective Enrollment Schools; and in New York City they are Screened or Specialized Schools.

we use public records to search for changes in admissions policies at each of these schools. Finally, we link these data with nationwide school-cohort-level records that capture shifts in demographic and academic performance over time.

Using this newly compiled dataset, we document geographic variation in the availability of ASHSs and patterns in their representativeness by socioeconomics and race. These selective schools are located across many parts of the US, usually in large urban school districts that serve a significant fraction of the US student population. We show that low-income, Black, and Latino students are underrepresented at ASHSs compared to their district composition. Meanwhile, ASHSs consistently overrepresent higher-income, Asian, and White students.

Next, we leverage the staggered rollout of race-neutral admissions reforms in our national dataset to estimate their demographic effects on ASHS entering cohorts. About 30% of schools in our analytic sample implemented such policies.[2] Policy implementation timing is likely not exogenous in our context because selective schools may change their admissions policies in response to new social pressures within their districts, or because of stagnation or deterioration in their diversity. Therefore, our preferred empirical approach accounts for school-specific trends prior to the policy change. Event study estimates show sharp shifts in demographic composition following policy adoption, implying that our estimates represent policy effects rather than the continuation of pre-existing trends.

Using a difference-in-differences (DID) approach that accounts for school-specific pre-trends, we find that, on average, race-neutral AA policies lead to a 14 percentage point (40%) increase in low-SES enrollment at selective high schools, compared to status quo ASHSs. We show that these policies also lead to a 6.5 percentage point (28%) increase in URM enrollment shares, which is driven by a 4 percentage point (12%) decline in the enrollment share of White students and a 2.4 percentage point decline in the enrollment shares of other race students. On average, these policy changes have little to no impact on Asian enrollment shares at selective schools.

Race-neutral AA policies that encourage diversity can vary in their structure and implementation. In our data, we identify four distinct race-neutral AA policies: Low-Income Quotas (LIQ) require a fraction of admitted/enrolled students be from low-SES backgrounds; class rank-based policies guarantee admissions to a portion of the best-performing students from each middle school within the district; geographic rank-based admissions prioritize admissions based on a student's residential location; and Submarginal Conditional Acceptance (SCA) admits low-SES students who score marginally below admissions cutoffs and successfully attend a bridge program between 8th and 9th grade.

[2]We restrict our analysis to policies that were not replacing race-based affirmative action. For example, we exclude ASHSs from Chicago, where a race-neutral policy replaced a race-based policy.

We find evidence of differential impacts between these different versions of race-neutral AA admissions. All four admissions policy types lead to increases in free lunch and URM enrollment shares. However, LIQs and Geographic Ranking admissions policies lead to significantly larger impacts on FL enrollment shares than class ranking and SCA policies. This difference may be due to Class Rank and SCA policies having a narrower scope by targeting only students at the top of their eighth grade class or scoring slightly below the cutoff, while LIQs and Geographic Ranking admissions policies have wider scopes in their admissions requirements to accept low-income students in order to increase their representation.

Next, we provide evidence that race-neutral admissions policies have minimal negative effects on schools' average academic performance, with treated schools experiencing a 1.7% decline on their average ACT/SAT scores, or a 0.53 point decline in the average ACT score.[3] We find heterogeneous impacts on academic achievement by type of admissions policy type. Specifically, we find that SCA and Class Ranking admissions policies have negligible impacts, while LIQs lead to significant declines in average ACT/SAT performance. This difference is due to Class Rank and SCA policies being academically stricter by targeting only students at the top of their eighth grade class or scoring slightly below the cutoff, while LIQs allow for separate admissions cutoffs for students meeting non-academic criteria.

We investigate potential mechanisms that may help explain how race-neutral affirmative action policies lead to declines in average ACT/SAT scores among treated cohorts. First, we provide evidence that the race-neutral AA policies negatively impact incoming cohorts' academic performance, which is in line with past research. However, these declines cannot fully explain the ACT/SAT effect. Moreover, we find suggestive evidence that the AA policies negatively impact social interactions based on measures of bullying, harassment, and suspensions, a mechanism that prior research has not examined.

To contextualize our estimates for ASHSs in districts with varied demographic composition, we next focus on the trade-offs between ASHSs' demographic diversity and average academic achievement. We estimate low-SES enrollment effects that are similar to predictions from counterfactual-based studies, but find smaller URM and academic effects than counterfactuals suggest. We additionally estimate similar trade-offs when we estimate a linear regression using the simulation-based policy impact predictions as data points.

Lastly, our estimates of the impact of changes in FL or URM enrollment shares on test scores allow us to consider how ASHSs' average test scores would change if they implemented admissions policies that closed ASHSs' FL or URM gaps relative to their local school districts.

[3]In this study, we rely on ACT and SAT scores as our primary indicator of students' academic performance. However, it is important to acknowledge that these standardized assessments may not fully reflect academic ability, as students from more affluent backgrounds often benefit from greater access to test preparation services and other supports that can boost performance (e.g., Goodman et al. (2020)).

We show that closing these gaps would, on average, lead to ACT/SAT performance declines at treated ASHSs of 2.85 percentile points, keeping them near the top of the ACT/SAT distribution.

Relationship to the literature. We make important contributions to several strands of research. First, we contribute to a growing literature that examines the potential impacts of race-based and race-neutral admissions policies on the socioeconomic and racial diversity of academically selective pre-collegiate schools. This question is increasingly important as educational institutions around the country are shifting away from purely achievement-based or race-based admissions to more race-neutral policies that prioritize low-SES students.

Despite this trend, research into the potential impacts of different admissions policies on selective K–12 school diversity is limited to theoretical and simulation-based studies (Cestau et al., 2017; Corcoran & Baker-Smith, 2018; Ellison & Pathak, 2021; Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022). These studies take the observed applicant pool and their application metrics as fixed, then re-run the admissions process under alternative rules to simulate admitted cohorts. The resulting freshman cohorts are then compared to the actual or alternative simulated cohorts.

Counterfactual simulations implicitly assume that all decision-making occurs on the institutional side: applicant behavior, application strength, and school preferences are held constant across policies. Yet evidence from higher education shows that students may respond strategically to admissions rules by adjusting test scores or class rank (Akhtari et al., 2024; Cullen et al., 2013). Simulations also assume homogeneous responses across demographic groups, despite research showing that families of different racial and socioeconomic backgrounds engage with school choice systems at different rates (Barrow, 2002; Crespin, 2022; Lawrence & Mollborn, 2017; Posey-Maddox et al., 2025). Thus, while counterfactual simulations offer valuable theoretical benchmarks, they rely on strong assumptions about family behavior and admissions dynamics. We test these hypotheses using real-world policy variation to study the impacts of race-neutral AA on incoming cohorts' demographics and average academic performance.[4]

We further contribute to this literature, which focuses on ASHSs within single school districts. More specifically, they focus on two of Chicago's Selective Enrollment High Schools (Ellison & Pathak, 2021), New York City's Specialized High Schools (Corcoran & Baker-Smith, 2018), and Boston's three Exam High Schools (Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022). We provide the first nationwide study of admissions policies. Among

[4]Furthermore, we contribute to this literature by examining the trade-offs between FL or URM enrollment changes and students' academic achievement in a standardized way, in order to benchmark with simulation-based predictions from the literature.

ASHSs in 47 school districts and 23 states, we observe 60 control ASHSs in 42 school districts and 28 treated schools within six districts.[5] This enables us to examine differential impacts from alternative AA admissions policies, allowing for variation in school and district-level implementation.

We also contribute to a growing literature on how changes in school composition affect social behavior and intergroup relations. This literature emphasizes that increased exposure to demographic diversity in school can improve individuals' social tolerance and reduce prejudice. Recently, Billings et al. (2021) studied a natural experiment where the removal of a desegregation plan led to changes in elementary and middle school diversity and found that increased exposure to minorities led White students to be less "racially conservative" during adulthood. Similarly, Carrell et al. (2019) find that White students who are randomly assigned to a Black roommate in their freshman year of college are more likely to subsequently choose a Black roommate. Outside the United States, Rao (2019) finds that increased socioeconomic diversity in preschool admissions to elite Indian schools generates short-run increases in tolerance and out-group friendship during second and third grade. In contrast, in the context of race-neutral affirmative action at academically selective high schools, we find increases in suspensions and race-based discrimination.

We argue that this contrast reflects differences in institutional context rather than conflicting evidence about intergroup contact. Prior empirical work focuses on settings without active affirmative action in admissions (Billings et al., 2021; Carrell et al., 2019) or on policies implemented at very young ages, when students may be less aware of policy dynamics (Rao, 2019). Our setting differs along several dimensions. We examine the activation of affirmative action in high-stakes admissions to academically selective high schools, where students are acutely aware of the policy change and where notions of merit and competition are salient. These schools feature limited excess capacity and sharp status hierarchies, which may heighten competition and grievances among incumbent or marginally admitted students. In this environment, increased diversity may generate social frictions that manifest in disciplinary outcomes rather than the prosocial behaviors documented in prior research. Our findings therefore suggest that the social consequences of intergroup contact depend on institutional context.

Lastly, our study complements a large and growing literature that examines access-oriented admissions policies on US universities' economic and racial composition. Those studies focus on the impacts of Top Percent and Holistic Review admissions on lower-income and URM undergraduate enrollment in California and Texas (Black et al., 2015; Bleemer, 2023; Cortes, 2010; Kapor, 2020; Klasik & Cortes, 2022). A key difference between our

[5]For this statistic, we count all NYC school districts as a single district.

context and the higher education literature is that we focus on admissions policy changes at schools that do not use race-based policies before the policy change. This way, we estimate a clearer policy impact from the introduction of race-neutral affirmative action rather than the impact of jointly removing race-based affirmative action and introducing race-neutral affirmative action.

Roadmap. This paper is organized as follows: Section 1 discusses the institutional and policy context. Section 2 describes the data collection process and analytic sample. Section 3 describes the empirical approach. Sections 4 and 5 present the main results, while Section 6 examines potential mechanisms. Section 7 benchmarks our results and examines the trade-offs between ASHSs' low-income or URM enrollment and average academic achievement. Finally, Section 8 concludes.

1 Institutional Background

1.1 Where and what are academically selective public high schools?

ASHSs, a subset of magnet/choice schools, typically offer advanced courses and extracurriculars that are not available in conventional high schools (Finn & Hockett, 2012; Hoxby, 2003). Their rigorous preparation often leads elite colleges to view ASHSs as dependable "feeders", making ASHS admissions a potentially important first step for elite college admissions (Hoxby & Avery, 2013). Moreover, these schools frequently play a significant role in shaping families' school choices, with over 60% of eighth graders applying for admissions in some districts (Feerick Center for Social Justice, 2019; Perez, 2018).

ASHSs are viewed by local politicians and school district administrators as a source of pride and a way to attract and maintain engagement from middle- and upper-middle-class families with urban public school systems (Finn & Hockett, 2012; Hoxby, 2003). Figure 1 shows that these schools exist in many regions of the United States. However, consistent with the idea that they are intended to generate and maintain public school enrollment from higher-income families in lower-income areas, ASHSs are mostly concentrated within large, urban districts.

While all schools included in this study are academically selective in some way, their methods for admitting incoming cohorts vary widely. This is exemplified by two examples from different sides of the selection spectrum. Stuyvesant High School in Manhattan is a famous example of a school that takes the exam school moniker literally in its admissions system—basing admissions solely on an exam score (Stuyvesant High School, 2023). Stuyvesant is part of a subset of New York City public schools that are called Specialized

High Schools. The New York City Specialized High Schools admission test (SHSAT) is offered to 8th and 9th grade students each year, and students are placed in rank order based on their scores on the exam (Dobbie & Fryer, 2014). Given the distribution of scores and the number of seats in the entering class, specialized high schools set a threshold score for admissions. Students also rank the schools that they wish to attend, so if they beat the cutoff score for their first-choice school, they are granted admission. If not, they drop to their second choice which may have a lower threshold score, and so on.

In other districts, the current admissions system eschews tests altogether, relying on a slew of other methods for evaluating academic performance and potential. For example, no standardized test scores are collected during the admissions process for any of the ASHSs in Washington, D.C. (DCPS, 2023). Instead, each school relies on its own mix of grade cutoffs, essays, recommendation letters, and interviews to determine its admissions offers.

1.2 Admissions Policy Changes

For decades, ASHSs across the US have faced criticism for not being socioeconomically or racially/ethnically diverse.[6] Lowell High School in San Francisco, California, provides a useful case study for the controversies and forces that drive ASHSs to alter their approach to admissions. Lowell became an ASHS in 1966 after it became oversubscribed, screening students based on their GPA. In 1983, the consent decree from San Francisco NAACP v. San Francisco Unified School District mandated that no single race or ethnic group could make up more than 40 percent of enrollment.[7] By 1995, this policy led to Chinese American students being held to higher standards on the admissions test than other groups (Woo, 1995). In February 1996, the San Francisco Board of Education changed Lowell's policy to fill between 70-80 percent of admission slots based only on grades and test scores. The remaining 20-30 percent were then filled with consideration for socioeconomic status, extracurriculars, past challenges, and race (Woo, 1996). In 2001, following the settlement of Ho v. San Francisco Unified School District which disallowed race-based admissions decisions[8], Lowell first implemented a three-level admissions system that closely resembles the current policy.[9] For the 2024-25 school year, Lowell's admissions system admits students in the same three bands.[10] First, 70 percent of seats are allocated to students based on middle school grades and score on the admissions test. Another 15 percent of seats are distributed to applicants who are evaluated based on a rubric that includes various indicators of economic and personal

[6] We provide more details about representation in Section 2.5.
[7] https://law.justia.com/cases/federal/district-courts/FSupp/576/34/2593333/
[8] https://clearinghouse.net/case/9799/
[9] https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=AZS28M7BDEA9
[10] https://drive.google.com/file/d/1SGCkuFxJvcVIdaYJK8MWpvo1pLrfvyYT/view

hardship, extracurricular pursuits, and service (students must still meet minimum grade and admissions test score requirements). The final 15 percent of seats are allocated to students who attend middle schools that are underrepresented at Lowell among the previous three Lowell cohorts (students must still meet a minimum score and must be nominated by their principal).[11]

ASHSs' admissions vary by context. In this paper, we focus on four race-neutral affirmative action admissions policies that aim to diversify the demographics of ASHSs. First are Low-Income Quota admissions policies that prioritize students who are categorized as low-income (New York City). Second are Eligibility in the Local Context (ELC) policies, which allocate a minimum number of seats for the top-performing students from each middle school in the district (Alexandra, VA; Charleston, SC; Boston, MA; New York City; San Francisco, CA). Third are Priority Group policies that target students who meet certain special circumstances, such as low parental educational attainment, living in a low-income neighborhood, or attending a low-income middle school (Austin, TX). And fourth is a Submarginal Conditional Acceptance (SCA) policy, referred to as the Discovery Program in New York City, which allows for an alternative route to admission for disadvantaged students in the city's specialized high schools through an invitation-only summer training program. We provide a detailed description of each of these policies in Appendix B.

2 Data

This study links four data sets to investigate the impacts of admissions policy changes at ASHSs. First, we use a database of ASHSs across the United States. Second, we create a database of ASHS admissions policy changes from 2010–2023. Finally, we use school enrollment data and standardized test score data to measure outcomes.

2.1 Database of Academically Selective High Schools

In defining ASHSs, we follow Finn and Hockett (2012) and require six conditions. Each school must be public, have a 12th grade, be a stand-alone school, teach a college-prep curriculum, select students on academic criteria, and have competitive admissions. Finn and Hockett (2012) identify 165 ASHSs in the United States. We identified 13 additional ASHSs to include in our sample through historical lists of ASHSs and by reviewing admissions policies for the top 150 high schools in the U.S. News & World Report rankings (2023).[12] We drop

[11] The full set of observed policy changes is described in Section 2.

[12] Although our process of identifying ASHSs may not be exhaustive, this would not affect the internal validity of our main identification strategy.

six schools from Finn and Hockett (2012) that are not found in the National Center for Education Statistics Common Core of Data, leaving us with a list of 172 ASHSs.

2.2 Admissions Policy Changes

We create a new panel data set of 122 admissions policy changes affecting 76 ASHSs between the late 2000s and early 2020s. We use official district and school websites to ascertain current school admissions policies. From there, we check for archived historical policies using the Internet Archive Wayback Machine. When available, we rely on these official admissions materials. Policy changes are recorded when there are observable adjustments to the application requirements or rubric for evaluating applications. Alternatively, when an ASHS's admissions policy documents are not regularly and thoroughly archived online, we search local news publications for mentions of the ASHS and descriptions of its admissions policy.

We identify 56 instances of race-neutral admissions policy changes related to student characteristics during our sample window from 2010–2023.[13] We denote schools as potential controls if they have no observed policy change of any type between 2010 and 2023.

2.3 School Demographics Data

We link our admissions policy data with the National Center for Education Statistics' Common Core of Data (CCD), a comprehensive database of public school enrollment by free-lunch eligibility (FL) and race/ethnicity (U.S. Department of Education, 2023).[14] The CCD provides race/ethnicity enrollment counts at the grade-by-school-by-year-level, which allows us to study changes at the 9th grade cohort-level. On the other hand, the free-lunch enrollment counts data are only available at the school-by-year-level, limiting us to study socioeconomic compositional changes over time at the school-level.[15] We impute the average between the previous and following school-year information when a school has missing demographic enrollment information. From this data, we restrict the sample for analysis to include only selective schools classified as traditional schools rather than alternative schools, vocational schools, or college-based programs. This data allows us to calculate enrollment shares by socioeconomic status and race/ethnicity.

[13] These exclude policy changes based on replacing application components (e.g. which specific exam is used).

[14] As suggested by the NCES, we calculate free-lunch based on the non-missing maximum between free-lunch eligible and direct certification counts (Sinclair & Chen, 2020).

[15] Through various initiatives, some schools have instituted universal school lunch programs over time, however Section 4.4 shows that our main results are robust to excluding observations and/or schools that report having 100% free-lunch eligibility.

2.4 Standardized Test Score Data

We also compiled a panel dataset of school-level average ACT and/or SAT scores, a proxy for college readiness. To do this, we collected publicly available data published by state departments of education or individual school districts. We relied on the Internet Archive Wayback Machine or FOIA requests when records were not directly available from official sources.

We map composite SAT scores into equivalent composite ACT scores with published concordance tables.[16] We use the ACT scoring scale because it remained consistent over time, while the SAT changed its scoring system in 2016. When both raw ACT and SAT scores are available, we attempt to analyze the primary exam for a given school and year. First, we use the exam with the most test takers for each school-by-year observation when this number is available. For schools that do not report the number of test takers, we prioritize based on score availability over time–we use the score from the exam with the least missing data. Next, when schools make clear switches between ACT and SAT as their primary exam, we allow their primary exam to change over time. We also interpolate average scores when there is a single year of missing data for the primary exam by taking the average of the one-year lead and lag. We use ASHSs' average ACT/SAT scores to estimate academic performance for the corresponding year's cohort of graduating seniors.[17]

We require a four-year lag to observe the testing outcome from an admissions policy change, because the reported average scores are based on the year that a cohort graduates from high school. This cuts our analytic sample of ASHSs down by about 29%. Specifically, we only have standardized test score data for 46% and 82% of the treated and control schools from our final analytic sample, respectively. Therefore, we limit our academic performance analysis to the sub-sample for which we have both demographic and ACT/SAT data, while reporting demographic impacts for both samples (with and without ACT/SAT data).

[16] For scores between 2005 and 2016 we use the concordance table found at: https://blog.prepscholar.com/act-to-sat-conversion. For scores starting in 2017, we use the concordance table found at: https://satsuite.collegeboard.org/media/pdf/guide-2018-act-sat-concordance.pdf

[17] Before the 2016-17 school year, both ACT and SAT score reporting was based on the graduating class. For example, this means that fall 2013 entrants are assigned to the average spring 2017 test scores. Beginning with the 2016-17 school year, there was a change in SAT score reporting such that, instead of reporting scores for the graduating class, averages were based on the test takers in a given year. We assume that most ASHS test takers are likely in their junior year rather than their senior year. Therefore, for cohorts starting before fall 2014 in schools where we prioritize the SAT instead of the ACT, we assign them spring test scores from three years later (their junior year). For example, for cohorts entering fall 2014 and later, we assign the SAT scores from their junior year (e.g. fall 2016 entrants are assigned to the spring 2019 average SAT scores).

2.5 Descriptive Statistics

2.5.1 Demographic Representation at ASHSs

Our compiled dataset allows us to illustrate the socioeconomic and racial representation at ASHSs across the US. Figure 1 shows that ASHSs are located in many parts of the United States, but are heavily concentrated in some of the largest urban districts in the country. Therefore, to appropriately describe the under- or over-representation of different demographic groups, we compare the enrollment share of a demographic group at each ASHS relative to the enrollment share of that group in the school district where the ASHS is located.

ASHSs underrepresent economically disadvantaged students. Panel A in Figure 2 shows the gap between each ASHS's free-lunch enrollment share (start of the arrow) and the enrollment share of that group in that ASHS's district (end of the arrow). In our dataset, 72% of ASHSs would have to increase their FL enrollment shares to reach full representation, with an average gap of -18 percentage points. Similarly, panel B shows that ASHSs also underrepresent Black and Latino students, with an average gap of -20pp, 76% of ASHSs would need to increase their URM enrollment shares to match their districts.

On the other hand, Asian and White students are overrepresented at ASHSs. Panel A in Figure 3 shows that 64% of ASHSs have a larger White enrollment share than their district, with an average representation gap of 9pp. Similarly, Panel B shows that 75% of ASHS have an overrepresentation of Asian students, with an average gap of 10pp.

2.5.2 Analytic Sample of Exam Schools

Our main goal in this paper is to compare treated ASHSs to ASHSs that did not implement changes to their admissions policies during our sample period. Table 1 compares our outcomes of interest across several relevant school groups during the 2015-16 school year—prior to any of the policy treatments that meet the criteria for inclusion in our analytic sample.[18]

Column (1) of Panel A describes ASHSs that appear in the NCES Common Core of Data, have non-missing enrollment-by-grade-by-race data in 2015, are classified as traditional high schools, and have a 9th grade class.[19] Columns (2)-(3) of Table 1 describe the main analytic sample, which excludes observations with missing race or FL enrollment data.

More specifically, Column (2) averages across schools in our sample that were never treated by an admissions policy change during our sample period (2010 through 2023). Column (3) restricts to those schools that were treated by a race-neutral admissions policy

[18] We discuss our decision rules for simultaneous or consecutive policy changes in appendix C.

[19] Other possible classifications include special education, vocational, alternative, and reportable programs.

change that targeted prospective students based on their characteristics.[20] We also restrict the treated schools to require that there are at least six years without a policy change before the relevant policy change goes into effect. This leaves our final sample of 60 control and 28 treated ASHSs. We utilize an unbalanced panel in which schools are treated for anywhere between 1 and 4 periods depending on the existence of subsequent policy changes and data availability, due to policy recency—Table A2 lists the treated schools in our analysis sample. Notably, the majority of our treated ASHSs are from New York City.[21] Untreated and SES treated schools have similar shares of FL and White students. However, compared to their counterparts, treated schools have a higher proportion of Asian students—36% compared to 16%, on average—while control schools have a higher share of URM students—51% compared to 23%.

Panel B of Table 1 presents summary statistics for the sub-sample of ASHSs for which we can compile ACT/SAT data. Columns (2) and (3) show that the treated schools, on average, have an ACT score of 31, while the control schools have an average score of 24. In Table A3, we show that the sub-sample with ACT/SAT data exhibits similar demographic patterns to those of the main sample.

3 Empirical Strategy

We estimate the effects of implementing race-neutral affirmative action on the demographic composition and academic achievement of academically selective high schools. Our approach compares changes in the differences in outcomes between ASHSs that implement race-neutral affirmative action admissions and ASHSs that continue with their status quo admissions during the sample period. To identify these effects, we leverage geographic and temporal variation in a staggered difference-in-difference (DID) approach.

More specifically, we estimate regressions of the following form:

$$Y_{sdt} = \alpha + \beta * AdmissionsChange_{st} * Post + \theta_t + \mu_s + \epsilon_{sdt} \tag{1}$$

where Y_{sdt} represents the outcome of interest at school s in district d in year t. Depending on the analysis, the outcome of interest is either the school's enrollment share for a specific demographic group (FL eligible, White, Asian, or URM) or average ACT/SAT score. Our key variable of interest, $AdmissionsChange_{st}$, is a measure of treatment dosage in the corresponding year. Our race enrollment shares and ACT/SAT data are available at

[20] Twenty-four schools from Column (1) are excluded from Columns (2) or (3) because they had admissions policy changes during the sample period that were not directly related to applicants' SES.

[21] We show that our policy impact estimates are consistent with and without treated NYC ASHSs.

the cohort-by-school level, capturing immediate policy impacts on each incoming cohort ($AdmissionsChange_{st} = 1$ for treated schools when analyzing race- and test-based outcomes). On the other hand, our FL-eligibility data is only available at the school-by-year level. This means that in the year of the policy shock, only one out of four cohorts in a high school is treated ($AdmissionsChange_{st} = 0.25$); by the fourth year, all cohorts in the school-by-year observation are treated ($AdmissionsChange_{st} = 1$). [22] We set $AdmissionsChange_{st}$ to 1 for all treated units during the pre-policy periods. Our control group of untreated schools is "uncontaminated" in two main ways. First, in the sense that it only includes ASHSs that did not experience any policy change in our sample period ($AdmissionsChange_{st} = 0$ for every year in the sample).[23] Second, it excludes non-ASHSs that may have experienced spillover-effects, where a reform could have pushed students from ASHSs into non-ASHSs, and vice-versa.[24]

This model includes year fixed effects, θ_t, to control for any time-varying changes in school demographics that are experienced across schools. We also include school fixed effects, μ_s, which control for time-invariant quality differences between schools. The school fixed effects account for any underlying school characteristics, such as prestige or desirability, that are known by school applicants before a policy change. The inclusion of both year and school fixed effects means the estimates are identified from within-school changes in demographics or academic performance after an admissions policy change.

The coefficient of interest, β, represents the DID estimate of the impact of a full-dose policy change on the outcome of interest after implementation, relative to the outcome prior to the policy change. Note that the model excludes $AdmissionsChange_{st}$ without the post indicator because, like other time-invariant school characteristics, this is absorbed by the school fixed effects. Standard errors are clustered at the district-level to account for similarity of policies and treatment timing across ASHSs within a district.

Policy implementation may not be exogenous. For example, ASHSs may decide to change their admission policies because of stagnation or deterioration in their diversity, or because of new social pressures within the district. To ensure we avoid potential bias in our estimates due to school-level variation in pre-implementation trends, we follow the literature and di-

[22] In robustness section D, we further discuss the implications and benefits of using a dose-response approach in our setting.

[23] We allow ASHSs that had multiple admissions policy changes, as long as they meet our requirement that the treatment timing does not overlap with a previous or future admissions policy change at the same school. That is, we require that the policy change must be at least six years after any previous policy change and at least one year prior to any future policy change. In practice, no school appears more than once in our analysis sample, though many schools have multiple policy changes.

[24] In Section 4.4, we show that our results are robust to excluding control ASHSs within the same district as treated ASHSs that may have experienced spillover effects.

rectly control for pre-trends by applying a two-step de-trending method (East et al., 2023; Goodman-Bacon, 2021).[25] Specifically, we estimate linear outcome pre-trends for each school using only data from before the policy change, which we extrapolate for each school through the post-shock period. Next, we subtract the fitted values from the observed values for each outcome, removing any differential pre-trends. This does not change the estimated trend breaks; it only modifies the orientation of the event-study coefficients (Goodman-Bacon, 2021).

Two main assumptions underlie our DID approach. First, any differential pre-trends must be approximately linear to be accounted for by our specification. This assumption is analogous to parallel pre-trends in the traditional DID specification. We test this assumption by conducting event study analyses that explore the yearly policy change effect on the outcomes of interest over time, after removing school-specific linear pre-trends. These event studies provide evidence on whether there are any remaining nonlinear differences across ASHSs in the pre-period.

We estimate a standard event study using regressions of the following form:

$$Y_{sdt} = \lambda + \sum_{\substack{\tau=-6,\\ \tau\neq -1}}^{3} \Big(\gamma_\tau * AdmissionsChange_{st} * 1(RelativeYear = \tau)\Big) + \theta_t + \mu_s + \epsilon_{sdt} \quad (2)$$

where the γ_τ coefficients represent the change in the relationship between an eventually/fully treated unit of observation and the outcome of interest in relative period τ compared to the relationship in relative period -1, which is the reference period.

Recent econometric studies highlight potential issues with the two-way fixed effects (TWFE) model and its corresponding event studies (Roth et al., 2023). Sun and Abraham (2021) highlight that the traditional TWFE event study yields consistent results only if certain stringent assumptions about homogeneous treatment effects across treated groups and/or across time are met. Therefore, to address concerns about the robustness of our main results and to allow for heterogeneity in treatment effects across time and treated units, we present event studies that employ four distinct robust estimators introduced by Callaway and Sant'Anna (2021), Sun and Abraham (2021), Wooldridge (2021), and de Chaisemartin and D'Haultfœuille (2020). With slightly different approaches and assumptions, each of these estimators delivers consistent estimates even in the presence of heterogeneous treatment effects across time and/or treated units, by shutting down DID comparisons between newly

[25] In Section 4.4, we show that not controlling for pre-trends leads to negatively biased estimates with qualitatively similar patterns.

treated and already treated units.[26]

The DID approach also assumes that there are no contemporaneous unobservable shocks that are correlated with the treatment and the outcome of interest. Although this assumption is not directly testable, our catalog of ASHS admissions policy changes allows us to exclude schools with multiple simultaneous admissions adjustments from our analysis.[27]

4 Impacts on Demographic Enrollment Shares

4.1 Impacts on Enrollment Shares, by Free-Lunch Status

Figure 4 presents an event study based on equation 2, which estimates the dynamic relationship between the activation of an admissions policy and FL enrollment shares at ASHSs from five years before a policy change and up to three years after implementation. Each dot represents the estimated coefficient, and the bands represent the 95% confidence interval. The dashed vertical line separates the pre- and post-policy shock periods. Because the outcome of interest is a measure between 0 and 1, each γ_τ represents the percentage point difference in the relationship between the eventual or implemented admissions policy change and the outcome of interest in period τ, relative to when $\tau = -1$.

The event study shows that after accounting for differences in school-specific linear pre-trends the estimates of the changes in the relationship between treatment and FL enrollment shares are close to zero and relatively flat during the pre-shock period. Thus, the figure provides supporting evidence for the linear pre-trends assumption, suggesting negligible remaining non-linear differences across treatment and control ASHSs in the pre-period. The implementation of a race-neutral admissions policy change in period 0 is associated with a visually clear and significant increase in FL enrollment shares. The abrupt increase makes it clear that our estimates represent policy effects rather than the continuation of a pre-existing trend.[28]

Panel A in Table 2 presents analogous DID estimates based on equation 1 for our overall sample. Column (1) show that controlling for school-specific pre-trends, there is an average increase of 13.5 percentage points (pp) in FL enrollment shares in response to a full dose

[26] See Roth et al. (2023) for a more thorough description of the differences between these four methods.

[27] We discuss cases and our decision rules for simultaneous/consecutive policy changes in appendix C. For example, our sample of treated schools excludes all Chicago ASHSs because the policy change that took place in 2009 replaced a race-based AA policy with an SES place-based policy. While we would be interested in the impacts of the SES place-based policy in Chicago, the termination of the race-based admissions policy would contemporaneously impact admissions in a way that would obscure the impact of the SES-based policy.

[28] As a robustness check, in Section 4.4, we show that event studies based on robust estimators that allow for heterogeneity in treatment effects across time and treated units have very similar dynamics to those estimated by our main approach.

of treatment, which is statistically significant at the 5% level. Taking this estimate at face value, it suggests a 40% increase in FL enrollment shares. This finding shows that the admissions policies operate as intended, increasing the representation of the lowest-income students at ASHSs.

4.2 Impacts on Enrollment Shares, by Race/Ethnicity

The evidence above strongly suggests that the implementation of race-neutral affirmative action policies had the first-order effect of improving the representation of low-income students, as proxied by FL eligibility. While these policies were intended to be race-neutral, expanding access for low-income students may have differential impacts by race due to the widely documented strong correlation between family income and race/ethnicity (Chetty et al., 2020).

Panel (a) of Figure 5 presents event study estimates for 9th grade cohorts' URM enrollment shares before and after race-neutral AA implementation, defined as the sum of Black and Hispanic enrollment shares. Accounting for differences in linear pre-trends, there are no remaining differential pre-trends between treatment and control ASHSs. In treated ASHSs, the period following the policy change exhibits a clear break in trend, with an immediate and statistically significant increase in URM enrollment shares within the first entering 9th grade cohort that is admitted under the new admissions policy.

Panels (b) through (d) in Figure 5 present event studies of the estimates for enrollment shares in 9th grade cohorts over time, by non-URM groups. Panel (b) shows that the increase in URM enrollment shares, discussed above, is almost entirely driven by an abrupt and significant decline in White enrollment shares. Importantly, panels (c) and (d) show no significant change in Asian and other race enrollment shares.

Table 2 summarizes the corresponding DID results based on equation 1. Column (2) in Panel A shows that, on average, implementing a race-neutral admissions policy leads to a 6.5 percentage point increase in URM enrollment shares (statistically significant at the 1% level), which is a 28% increase relative to the mean. On the other hand, there is a 4 percentage point (12%) decline in White enrollment shares, which is not statistically significant. Column (3) shows that, on average, there is also insufficient evidence to reject the null hypothesis that the race-neutral admissions policies have no impact on Asian enrollment shares.[29]

[29]The URM and White coefficients do not exactly counteract each other because there is a slight decline in other race enrollment shares of 2.4pp, which is not statistically significant.

4.3 Policy Heterogeneity

Table A5 displays the estimated policy impacts on FL enrollment shares (panel A), URM shares (panel B), White shares (panel C), Asian shares (panel D), and other race enrollment shares (panel E) for each of the four policy types (shown in each column).[30] Panel A shows that the Low-Income Quotas and Geographic Ranking policies had large and statistically significant impacts on FL enrollment shares, while middle school class ranking and the SCA policies had less half as large of impacts and were not statistically significant. This difference may be due to Class Rank and SCA policies having a narrower scope by targeting only students at the top of their eighth grade class or scoring slightly below the cutoff, while the other two policies have a wider scope, creating new admissions cutoffs for students meeting SES criteria.

Panel B of Table A5 shows that the estimated impacts on URM enrollment shares from each of the four policy types are not statistically different from each other. Furthermore, Table A5 shows suggestive evidence that the Low-Income Quotas and Geographic Ranking admissions policies' positive impacts on URM enrollment shares are driven by declines in White enrollment shares. However, we find suggestive evidence that the Class Rank and SCA policies' impacts on URM may be driven by declines in Asian and other race enrollment shares.

Our estimates notably differ from smaller policy impacts observed at the college level using natural experiments. For example, Kapor (2020) estimates a 5.4% effect from the Texas' Top Percent admissions policy on enrollment of students from high-poverty high schools, while Bleemer (2023) finds a negligible effect from California's Top Percent policy on low-SES student enrollment. Moreover, our estimates of the effects on URM enrollment are substantially larger than those documented in the college context (Bleemer, 2023; Cortes & Klasik, 2020; Kapor, 2020).[31] These contrasts underscore the importance of examining SES-based AA in selective high schools, where such policies seem to be more effective than similar initiatives in college admissions.[32]

[30] In appendix B, we provide more details on each of the SES-based admissions policies that were implemented by ASHSs in our analytic sample.

[31] One main limitation of these studies, is that they rely on the replacement of race-based AA with Top 10% policies. However, in our context we examine the transition from status-quo admissions that are not race-based, to race-neutral AA policies.

[32] In Section 7, we show that our estimated impacts on low-income and URM enrollment shares are more comparable to those from K-12 AA policy simulations.

4.4 Robustness

We now test the sensitivity of our estimated impacts on demographic composition by conducting a series of robustness checks. First, Table A4 shows very similar estimated effects when we limit the analysis to include all pre-shock periods but only the initial post-shock period to ensure a balanced sample.

Second, our FL estimates rely on a dose-response approach to generate coefficients that are interpreted in the same way for both socioeconomic outcomes and for race/ethnicity outcomes.[33] In Appendix D, we show that with proper scaling, our dose-response approach is mathematically equivalent to the traditional approach without dosage.

Third, Sun and Abraham (2021) highlight that the traditional TWFE event study yields consistent results only if certain stringent assumptions about homogeneous treatment effects are met. To allow for heterogeneity in treatment effects across time and treated units, we present event studies in Figure A1 that employ robust estimators introduced by Callaway and Sant'Anna (2021), Sun and Abraham (2021), Wooldridge (2021), and de Chaisemartin and D'Haultfœuille (2020). The figure suggests that the underlying assumptions are not strongly influencing our results and that our estimates are robust to bias from heterogeneous treatment effects across groups or time.[34]

A fourth potential concern is that, given the composition of our group of treated schools, our estimated effects are only driven by New York City schools. To test this, we replicate our analysis on two separate samples. First, we exclude all treated New York ASHSs. Then, we include only treated ASHSs from New York, with the same set of control schools used in both samples. Appendix table A6 presents quantitatively and qualitatively similar estimates of the impact of implementing SES-based AA on the low-income and racial composition at non-NYC ASHSs (in panel A) and at NYC-only ASHSs (in panel B).[35] This test suggests that the impacts we find are broadly representative of ASHSs, and not driven solely by NYC ASHSs.

Next, Table A7 shows that our results are robust to various sample restrictions. Column (1) shows the main results for comparison. To address the concern that perhaps our main results are driven by pandemic-era admissions policy changes, columns (2) and (3) show that we find similar estimates when we limit to pre-pandemic treated units and drop all obser-

[33] As described in section 2, our FL enrollment data is only available at the school-by-year-level, while the race enrollment data is available at the grade-by-school-by-year-level.

[34] The robust estimators cannot accommodate dose-response DID, therefore, the event study for FL enrollment shares in panel (a) is estimated in the traditional way, without dosage. As described in Appendix D, FL event studies that do not incorporate dosage will differ from our dose-response estimates in periods +0, +1, and +2, however, these results are equivalent with proper rescaling.

[35] Due to the smaller sample size, the estimates are noisier and have larger standard errors.

vations before 2020 or before 2021. Furthermore, through various initiatives, some schools, districts, and states have instituted universal free lunch programs, potentially impacting some of the schools in our sample. In columns (4) and (5) we estimate similar coefficients when we exclude all observations that report 100% FL eligibility or all ASHSs that ever report having a 100% FL eligibility.

Another potential concern is that in districts with multiple treated ASHSs, admissions policy changes in the most in-demand ASHS may have trickle-down effects for ASHSs that are less in-demand.[36] To address this concern, in column (6) of Table A7, we keep only the ASHS with the highest average ACT scores if there are multiple ASHSs in the district.[37] Additionally, in column (7) we drop all control ASHSs in districts with treated ASHSs from the sample to address the possibility that untreated ASHSs have their cohort composition impacted by the admissions policies at treated schools. In both subsamples, we estimate similar, although slightly larger, coefficients compared to our main sample.

Next, in Table A8 we estimate similar coefficients when we control for year-to-year district-level racial composition changes (column 2) or year-to-year district-level income composition changes (column 3). This provides evidence that we are estimating school-specific effects rather than district-wide demographic changes. Additionally, in Table A9 we show that using a stacked DID specification, as proposed in Wing et al. (2024), yields similar demographic results (shown in column 2). Lastly, in column (3), we show that failing to account for school-level enrollment pre-trends yields misleading smaller estimates of the policy impacts.

Another potential concern with our identification strategy is the existence of secular trends that differ systematically with treatment exposure. Our event studies suggest such trends are not biasing our estimates. In Appendix E, we conduct permutation tests as further evidence that our estimated demographic impacts are not identified based on secular trends. Although reversion to the mean may play some part in enrollment share changes, placebo tests show that our estimates are strongly linked to whether an admissions policy is implemented and to the specific timing with which the policy is implemented. Furthermore, if reversion to the mean is playing a role in our estimates, this further suggests that we are estimating a lower bound of the actual policy impact.

[36] If higher-ranked ASHSs admit more under-represented students, lower-ranked ASHSs may not be able to attract as many AA-targeted students, whether or not they also instituted a race-neutral AA admissions policy.

[37] Effectively, this drops non-Stuyvesant ASHSs in NYC and non-Boston Latin ASHSs in Boston.

5 Impacts on Average ACT/SAT Scores

Next, we examine how race-neutral AA admissions policies affect ASHSs' average academic performance, as proxied by performance on the ACT/SAT, at the cohort-level within schools. To do this, we focus on the sub-sample of ASHSs for which we have standardized test score data, as described in Section 2. Since the scores are based on cohorts' year of graduation, we require a four-year lead to observe outcomes with the correct timing. This reduces our sample by about 29%, leaving available standardized test score data for 46% of treated schools and 82% of control schools in the final analytic sample. Nonetheless, we provide supporting evidence for the generalizability of our results on test scores.

5.1 Results

Panel B in Table 2 shows that the test scores sample exhibits very similar demographic impacts to those from the main analytic sample, with an increase of 14.1 percentage points (42%) in FL shares and a 5.6 percentage point (35%) increase in URM enrollment shares, both of which are statistically significant at the 1% level. Again, the gains in URM enrollment shares come from declines in White and other race enrollment shares.[38]

The event study in Figure 6 shows that even with significant gains in FL and URM enrollment shares, the SES-based admissions policies have a small impact on the standardized test scores of the corresponding cohorts at treated ASHSs. Column (6) of Table 2 shows that the average ACT score for the treated cohorts decreased by 0.53 points, or 1.7%, which is not statistically significant. Using the 95% confidence interval, our estimate rules out an increase in average ACT scores greater than +0.11 points and rules out declines greater than -1.17 points.

Table A10 presents the estimated effects on test scores split by policy type.[39] Panel A shows that low-income quotas had the most negative effects on average test scores, with a decline of about 1.4 ACT points, or 5%. However, the other two policies have negligible effects on test scores that are not statistically significant. This difference may be driven by the fact that SCA and Class Ranking policies are stricter by targeting the top middle school students or those slightly below the overall admissions cutoff, rather than creating separate standards for students from low-income backgrounds.

[38] We present the corresponding event studies in Figure A2.

[39] We cannot separately estimate effects for geographic ranking admissions policies because, in our sample, these took place too recently to capture test score outcomes data.

5.2 Robustness

While our sample of treated schools with ACT/SAT data is smaller than our overall sample of ASHSs, we conduct robustness tests to examine the consistency of the results. Figure A3 shows that the pre-trends and post-shock effects are quite consistent across four robust estimators as well as the original two-way fixed effects specification, suggesting that the underlying assumptions are not strongly influencing our results.

Since five of the seven treated ASHSs with available test score data are in NYC, one may be concerned that the sample choice may be more impactful for the test score results. In Table A11, we re-estimate the test score regressions for separate samples of treated ASHSs in and outside NYC, again with the same set of control schools used in both samples. In columns (2) and (3), we find that the race-neutral AA policies led to declines in average ACT scores of 0.91 and 0.42, in treated non-NYC and NYC ASHSs, respectively. Therefore, our estimates remain consistent with our main findings that SES-based admissions policies have a small or negligible impact on ACT scores. Moreover, in column (4) we show that a stacked DID specification leads to a decline in average ACT score of 1 point, which is still within the 95% confidence interval of our main estimate.

6 Potential Mechanisms

In this section, we investigate potential mechanisms that may help explain how race-neutral affirmative action policies lead to declines in average ACT/SAT scores among treated cohorts. We consider two main channels. First, we analyze shifts in the academic preparedness of incoming cohorts, the primary channel emphasized by prior work. Second, we study changes in school social climate—a mechanism largely unexplored in the existing literature.

6.1 Incoming academic performance mechanism

A potential mechanism for the declines in average ACT/SAT scores could be that ASHSs are simply admitting lower-performing students. Indeed, simulation-based counterfactual studies consistently show that, on average, race-neutral AA admissions policies lead to declines in the incoming achievement of admitted students (Cestau et al., 2017; Corcoran & Baker-Smith, 2018; Ellison & Pathak, 2021; Rokkanen, 2015). In this subsection, we examine this potential mechanism using real-world policy variation.

Generally, AA is likely to displace only marginal students near the admissions threshold, not the highest achieving students (Hoxby, 2025). Therefore, to examine this potential mechanism we proxy for the incoming achievement of marginal students by examining the

performance and fraction of students taking the lowest-level high school math course, Algebra 1. ASHSs may have a substantial fraction of students taking advanced math courses during their freshman year of high school, but Algebra 1 would be the lowest-level math course offered in these academically selective schools.

To directly analyze the impacts of the AA policies on incoming academic performance, we use school-by-year-level Algebra 1 Regents exam data from New York City, which covers a large fraction of our treated sample of ASHSs.[40] We proxy for the number of students taking Algebra 1 by the number of students taking the Algebra 1 Regents exam in each school-by-year observation. Additionally, we proxy for incoming achievement by capturing the average score and the percent of students scoring in the highest performance category in the Algebra 1 Regents exam. For this analysis we must rely on only treated ASHSs, therefore, for the rest of this subsection, we employ the approach developed by Callaway and Sant'Anna (2021) for staggered DID with yet-to-be-treated units used as controls. Figure A4 replicates the ACT/SAT results using this yet-to-be-treated approach, showing that the AA policies lead to a decline on treated cohorts' ACT/SAT scores, across policy type.

Overall, we find evidence that while the AA policies lead to increases in Algebra 1 participation by low-SES students, incoming performance does not fully explain the declines in ACT/SAT scores. Figure A7 shows that there is an increase in low-SES students participating in Algebra 1, while not significantly impacting the participation rates of higher-SES students. This pattern is especially driven by the SCA policies, which have immediate significant impacts for low-SES participation in Algebra 1, without a change in the participation by higher-SES students.

Panel A of Figure A5 shows that the admissions change leads to a slow decline in Algebra 1 Regents score. However, these changes are not statistically significant. Panels (b) and (c) show that almost all of this effect is driven by the LIQ policies (about 2.5% decline), and no significant impact by the SCA policies (about a 1% decline). This is even though SCA policies led to large and significant increases on low-SES participation in Algebra 1. Furthermore, Figure A8 shows that the LIQ polices led to declines in the rate of students scoring in the highest Algebra 1 performance category for both lower- and higher-SES students, while the SCA policies did not change the rate of students scoring in the highest performance category.

In summary, these findings suggest that while the race-neutral AA polices led to an increase in Algebra 1 participation by low-SES students, these participation changes did not lead to significant performance changes in the Algebra 1 Regents Exam. This implies that

[40] We describe our Regents Exam data cleaning process in Appendix F.

students' incoming ability cannot fully explain the declines in average ACT/SAT scores.[41]

6.2 Social interactions mechanism

A second potential mechanism entails harassment, bullying, and discrimination. Affirmative action admissions can generate social tensions if students who benefit from these policies face hostility from peers who view themselves as more deserving of admission. As cohorts become more diverse, such dynamics may shape the school climate and, in turn, affect academic performance.[42]

To explore this potential channel, we draw on school-by-year-level measures of bullying, harassment, and related disciplinary incidents from New York State's Department of Education, New York City's Department of Education, and the U.S. Department of Education. Since the early 2010's, the NYS DOE has provided unredacted and uncensored counts of the number of incidents of harassment and/or discrimination. NYS (through 2017-18) and NYC (starting in 2018-19) break out these counts by the nature of the harassment (e.g. race-based).[43] We also obtain unredacted/uncensored data on the percent of students suspended in each school each year. Lastly, we obtain unredacted/uncensored data of race-based incidents, broken out by the number of reported cases (by victim's race/ethnicity) and the number of disciplinary actions (by perpetrator's race/ethnicity).[44] Combining these sources of data provide direct measures of peer conflict and potential intergroup tensions following the integration-focused policies.

Panel (a) in Figure A9 shows that the admissions policy shock leads to a statistically significant immediate increase in the fraction of suspended students, which is sustained through three post-shock years.[45] Furthermore, panel (f) show suggestive evidence that

[41] While our data allow us to examine incoming achievement being a potential channel through which ACT/SAT scores decline, we are unable to directly examine academic mismatch or whether students experience academic growth while attending an ASHS. Studies using regression discontinuity (RD) in Boston, Chicago, and New York suggest that ASHSs have negligible impacts on marginally admitted students' standardized test scores (Abdulkadiroglu et al., 2014; Angrist et al., 2023; Barrow et al., 2020; Dobbie & Fryer, 2014). However, emergent research suggests that income- or race-based AA that admits lower-achieving students may positively impact marginal students' high school test scores by providing them access to exam schools (Idoux, 2022; Rokkanen, 2015). Those studies focus on marginal students farther from the admissions cutoff or marginal students who face weaker high school alternatives in Boston.

[42] Research shows that exposure to disruptive environments can have negative effects on students' short- and long-run academic outcomes (Carrell & Hoekstra, 2010; Carrell et al., 2018).

[43] NYS provides unredacted and uncensored datasets, while NYC redacts all cells that have counts of one to five. We impute for redacted cells, by dividing the total harassment counts equally into each redacted non-zero cell.

[44] The CRDC provide data collected biennially and include the school years 2013-14, 2015-16, and 2017-18. We interpolate missing years linearly, taking the average of the one-year lead and lag.

[45] Panels (b) and (c) show that both LIQ and SCA policies exhibit the immediate impact, but only the SCA policies have a sustained increase in suspension rates.

SCA, but not LIQ, policies lead to an increase in race-based harassment/bullying incidents, which can help explain the increased suspension rate.[46] More specifically, Figure A10 shows suggestive evidence that SCA admissions policies led to an increase in disciplinary actions for race-related harassment or bullying by White student perpetrators (shown in panel a).[47] On the other hand, Asian and Hispanic students experience increases in race-related harassment or bullying (panels e and f).[48]

As described in Section 4, White students lost the largest enrollment shares from race-neutral AA policies. Therefore, the dynamics of White students harassing and bullying non-White students more once the AA policies are implemented may be explained by White students potentially viewing themselves as more deserving of admissions than their non-White peers. Perhaps these effects are driven by students' socioeconomic status, but our data on harassment and bullying does not capture this aspect of victims' or perpetrators' demographics.

7 Benchmarking

To contextualize our findings, we compare the magnitude of our race-neutral AA estimates to those reported by prior studies regarding race-based and race-neutral admissions policies.

To our knowledge, empirical evidence on the demographic and academic effects of AA policies in ASHSs is limited to simulation-based analyses. We summarize the existing studies in Table A1. To be included, a study had to: (a) use data, not solely theoretical models; (b) focus on academically selective K–12 schools; and (c) analyze affirmative action policies that assign admissions priority based on student characteristics (as opposed to modifying application requirements such as GPA or test scores).

We identified six studies that meet these criteria through targeted Google Scholar searches using terms such as "affirmative action admissions" and "exam school affirmative action." The studies examine different urban settings—including Boston, Chicago, and New York City—and cover applications to 1st, 7th, and 9th grade programs. Some rely on data from a single application cohort (Corcoran & Baker-Smith, 2018; Ellison & Pathak, 2021), while others use multi-year datasets spanning three to twelve cohorts (Cestau et al., 2017; Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022). Across contexts, the simulations eval-

[46] We find no such impacts on religion-, disability-, sex-, or sexual orientation-based harassment/bullying incidents.

[47] Our CRDC data only capture years where some SCA policies were implemented, but no LIQ policies had been implemented yet.

[48] Our data only allows us to make statements about average changes, as the data does not capture case-by-case incidents with direct victim and perpetrator characteristics.

uate a range of admissions policies, including: (a) no AA, (b) SES-based AA, (c) middle school class rank priority, (d) geographic-based rank priority, and (e) race-based AA. All of these studies report impacts on student composition (SES and/or race), while all except two examine impacts on average academic performance.[49]

Generally, these studies implement counterfactual admissions simulations by taking the actual applicant pool and their observed application components/metrics as a fixed snapshot and re-run the admissions process under alternative rules, keeping available seats constant. In practice, this means that applicants are ranked according to the policy-specific admissions criterion, accounting for quotas or admissions priorities, and offers are extended in descending order until all seats are filled. The resulting simulated cohort is then compared to the actually admitted or other policy simulated cohort.

Importantly, counterfactual simulations assume that all decision making is on the institutional side and that applicant behavior is static—who applies, their application metrics, and their preference rankings do not change when admissions rules change. However, research on the impacts of AA in higher education shows that students may behave strategically by improving their admissions test scores or class rank in response to admissions policy changes (Akhtari et al., 2024; Cullen et al., 2013). Furthermore, simulations assume that families from different demographic backgrounds would not respond differently. However, research suggests that families from different racial/economic backgrounds are not equally likely to engage with school choice policies (Barrow, 2002; Lawrence & Mollborn, 2017; Posey-Maddox et al., 2025). While these types of studies provide important theoretical insights, they must rely on assumptions about families' preferences and admissions dynamics. We test their hypotheses by employing real-world policy variation.

7.1 Policy Impacts

Impacts on Enrollment Shares, by Free-Lunch Status.– The forest plot in Figure A11 presents our estimated overall policy impact (red diamond) along with race-neutral (filled points) and race-based (hollow points) simulation-based predictions. We estimate a slightly larger policy impact on FL enrollment shares relative to the median simulated impact for race-neutral (dashed vertical line) and for race-based (dotted vertical line) admissions policies. However, these differences are not statistically significantly different from each other. The estimated

[49] To facilitate comparison across studies, we standardize the reported results, by expressing demographic impacts as percentage points changes and converting all achievement impacts into fractions of a standard deviation for whichever academic measure is used.

95% CI captures 26 out of 31 (83%) simulated predictions.[50]

Impacts on Enrollment Shares, by Race/Ethnicity.—The forest plot in Figure A12 shows that while our estimated overall policy impact on URM enrollment shares is positive and statistically significant, it is at the lower-end of the distribution of simulation-based predictions. Our estimate is significantly smaller than the median predictions from race-neutral and race-based AA simulations. This difference may be due to simulation studies relying on various assumptions, including not accounting for potential differences in families' responsiveness or social network constraints when navigating school choice systems (Barrow, 2002; Lawrence & Mollborn, 2017; Posey-Maddox et al., 2025).[51]

Impacts on Average ACT/SAT Scores.—When we convert our estimates to standard deviation units for comparison with the existing simulation-based literature, we estimate an overall policy impact of -0.10σ. The forest plot in Figure A15 shows that our estimated effect is less negative than the median impact for race-neutral and race-based admissions policies.[52] However, only the race-neutral median falls below our 95% confidence interval.

The forest plot in Figure A16 shows that the simulation-based median for each policy type is within the 95% CI of our corresponding estimated impact.[53] We estimate that the SCA policy has the smallest impact on test scores, but we do not have simulation-based predictions for this type of policy to compare to. Given differential impacts on FL and URM enrollment shares examined in Section 4, we present a more comparable academic impact metric that accounts for differential demographic changes in Section 7.2.

7.2 Trade-offs between Diversity and Test Scores

The policy impacts on test scores do not account for the differential impacts of each policy change on FL or URM enrollment shares. Therefore, we now focus on the trade-off between comparable changes in diversity and average academic performance.

We employ an instrumental variables (IV) approach to estimate the effects on average test scores from a change in FL or URM enrollment shares. Table A12 present 2SLS regression estimates of the trade-off between average test scores and demographic enrollment shares,

[50]The forest plot in Appendix Figure A13 shows that the simulation-based median for each policy type is within the 95% CI of our corresponding estimated impact on FL enrollment shares, with the exception of our LIQ estimate, which has a 95% CI that is larger than corresponding simulation-based median.

[51]The forest plot in Appendix Figure A14 shows that our estimated policy impacts on URM enrollment shares are significantly lower than the median of their corresponding simulated predictions.

[52]To make estimates comparable to each other, we convert all impacts to be relative to their standard deviations.

[53]More specifically, the 95% confidence interval for our low-income quotas estimate captures the estimates of nine out of 13 simulation-based predictions. Similarly, the confidence interval for our class-ranking policy estimate captures two out of three simulation predictions.

using the SES-based admissions policy shocks as the IV. Overall, the 2SLS estimate in panel A indicates that increasing FL enrollment shares by 10 percentage points decreases ASHSs' average ACT score by 0.96 point or 0.18σ. Similarly, Panel B shows that increasing URM shares by 10 percentage points decreases schools' average ACT score by 0.94 points, or 0.18σ.

The forest plot in Figure A17 shows that we estimate a smaller negative effect on test scores from an increase in FL shares than the majority of simulation-based predictions. However, Figure A18 shows that our estimate of the impact of a 10pp increase in URM shares on test scores is right at the median of the distribution of race-neutral simulation-based predictions.

An alternative approach to estimate the effect on average test scores from a change in FL or URM enrollment shares is to directly estimate a linear regression using the simulation-based predictions as data points. In Figure A19 we plot the raw, standardized effect of each policy on average test scores against the change in low-SES (left panel) or URM (right panel) enrollment shares caused by the same policy. The filled dots represent race-neutral simulations, while the hollow points are race-based policy simulations. For reference, we plot our overall 2SLS estimates with a red diamond. We overlay a linear regression on these data and cluster standard errors by study (we exclude our overall estimated impact from this calculation). This approach estimates that a 10 percentage point increase in low-SES or in URM shares leads to declines of 0.22σ or 0.17σ in average test scores, respectively. These estimates are in line with our estimated impacts of 0.18σ from Table A12.

7.2.1 How would closing representation gaps impact test scores?

As noted in Section 2.5, ASHSs often underrepresent free lunch eligible students, Black students, and Latino students while overrepresenting White and Asian students relative to the racial and socioeconomic composition of their districts. Our 2SLS estimates of the impact from changes in FL or URM enrollment shares on test scores allow us to consider how ASHSs' average test scores would change if they implemented race-neutral admissions policies that closed the FL or URM representation gaps relative to that of their local school district.

To do this, we first calculate the district-level share of underrepresented minority students and FL-eligible students for the 2015-2016 school year for every district with an untreated ASHS in our sample and for the year prior to treatment for every district with a treated ASHS in our sample. From these district-level shares, we calculate the representation gap for each ASHS by subtracting its enrollment shares from its districts' enrollment shares in the same year. Finally, we multiply the gap by the coefficients estimated in Panels B (for FL-eligible share) and C (for URM share) of Table A12 to generate the test score impact

from closing the FL or URM representation gaps.

Figure A20 plots the simulated effect of closing the FL representation gap on the average ACT score for each ASHS. Panel A shows that, on average, treated ASHSs' average ACT would decline from an average of 30.8 points to 28.5 points. This translates to an average decline of 0.85 percentile points, moving treated ASHSs from the 99.8th percentile to the 98.9th percentile (shown in Panel B).[54] More generally, ASHSs' ACT scores would decline by 1.8 points, which translates to moving the average ASHS from the 83.7th percentile to the 78.9th percentile, a drop of 4.8 percentile points.

Similarly, Figure A21 shows that closing the URM representation gap at treated ASHSs would lead to a decline of 2.2 ACT points. This translates to a decline of 2.8 percentile points, to an average ACT/SAT score at the 96.9th percentile.[55]

7.2.2 Heterogeneity by policy type

In panels B and C of Table A10, we examine heterogeneity in the trade-offs between impacts on test scores and impacts on FL or URM enrollment shares, by policy type. SCA and Class-Ranking policies have the smallest impacts on academic achievement, with estimated declines of 0.05σ to 0.14σ (neither statistically significant). On the other hand, Low-income quotas have the largest impacts on academic achievement, with statistically significant declines of 0.29σ to 0.31σ. One potential reason for SCA and Class-Ranking admissions policies having the least negative impacts on academic achievement is that both of these policies have narrower scopes and are more strictly beneficial to applicants based on quantitative academic measures, rather than Low-Income Quotas which may expand low-SES representation by setting separate admissions cutoffs.

The forest plots in Figures A22 and A23 show that our estimates by policy-type are generally consistent with the corresponding simulation-based predictions of the trade-off between diversity and academic achievement. However, our 2SLS approach generates large standard errors.

8 Conclusion

Selective educational institutions provide students with access to competitive learning environments, substantial resources, and high-achieving peers (Chetty et al., 2023; Dobbie &

[54] School average ACT percentiles are calculated based on each ASHS's percentile ranking relative to school-wide average ACT scores for high schools we were able to obtain ACT/SAT data for, which include AL, CA, DC, DE, FL, GA, IL, KS, KY, LA, MA, MO, NC, NV, NJ, NY, PA, SC, TN, TX, and WI.

[55] More generally, ASHSs would decline by 2.2 ACT points, dropping the average to the 80.1th percentile.

Fryer, 2014; Finn & Hockett, 2012). For decades, these institutions have struggled to diversify their campuses and expand access to high-quality education for students from different social backgrounds. While universities have been at the center of this debate, the topic is equally controversial for admissions to highly-coveted K–12 academically selective schools.

In this paper, we provide the first national, empirical evidence on the effects of race-neutral affirmative action policies at academically selective high schools (ASHSs). We construct a novel dataset of selective public high schools and their admissions policies since 2010 and link these data to national cohort-level demographic and test score records. Descriptively, we document substantial underrepresentation of low-income, Black, and Latino students at ASHSs relative to their districts, alongside overrepresentation of higher-income, Asian, and White students.

Leveraging geographic and temporal variation in admissions reforms, we show that race-neutral policies significantly increase low-income and underrepresented minority enrollment shares. We document meaningful heterogeneity across policy designs: policies that allow greater flexibility in admissions cutoffs produce larger diversity gains, while stricter systems generate more modest demographic changes.

Despite admitting students with lower prior achievement on average, race-neutral affirmative action policies lead to only modest declines in ACT/SAT performance, leaving treated schools near the top of the national achievement distribution. To better understand these effects, we examine potential mechanisms. We find that part of the decline in test scores reflects lower average incoming preparation among admitted cohorts, though this channel does not fully explain the overall achievement effects. We also provide suggestive evidence that reforms may affect the social condicions, with increases in reported bullying, harassment, and suspensions following policy adoption.

This paper fills a gap in the economics and education literature by providing the first empirical evidence of race-neutral AA using real world policy variation. Furthermore, our analysis provides insights for policymakers and administrators at selective educational institutions into the potential effects of various race-neutral affirmative action admissions policies which do not explicitly rely on race—making them permissible under the latest Supreme Court rulings against race-based affirmative action.

References

Abdulkadiroglu, A., Angrist, J. D., & Pathak, P. A. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, *82*(1), 137–196. https://doi.org/10.3982/ECTA10266

Akhtari, M., Bau, N., & Laliberté, J.-W. (2024). Affirmative Action and Precollege Human Capital. *American Economic Journal: Applied Economics*, *16*(1), 1–32. https://doi.org/10.1257/app.20210807

Angrist, J. D., Pathak, P. A., & Zarate, R. A. (2023). Choice and consequence: Assessing mismatch at Chicago exam schools. *Journal of Public Economics*, *223*, 104892. https://doi.org/10.1016/j.jpubeco.2023.104892

Arcidiacono, P., & Lovenheim, M. (2016). Affirmative Action and the Quality–Fit Trade-off. *Journal of Economic Literature*, *54*(1), 3–51. https://doi.org/10.1257/jel.54.1.3

Barrow, L. (2002). School choice through relocation: Evidence from the Washington, D.C. area. *Journal of Public Economics*, *86*(2), 155–189. https://doi.org/10.1016/S0047-2727(01)00141-4

Barrow, L., Sartain, L., & de la Torre, M. (2020). Increasing Access to Selective High Schools through Place-Based Affirmative Action: Unintended Consequences. *American Economic Journal: Applied Economics*, *12*(4), 29.

Billings, S. B., Chyn, E., & Haggag, K. (2021). The Long-Run Effects of School Racial Diversity on Political Identity. *American Economic Review: Insights*, *3*(3), 267–284. https://doi.org/10.1257/aeri.20200336

Black, S. E., Cortes, K. E., & Lincove, J. A. (2015). Academic Undermatching of High-Achieving Minority Students: Evidence from Race-Neutral and Holistic Admissions Policies. *American Economic Review*, *105*(5), 604–610. https://doi.org/10.1257/aer.p20151114

Bleemer, Z. (2023). Affirmative Action and its Race-Neutral Alternatives. *Journal of Public Economics*. https://doi.org/https://doi.org/10.1016/j.jpubeco.2023.104839

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The Long-Run Effects of Disruptive Peers. *American Economic Review*, *108*(11), 3377–3415. https://doi.org/10.1257/aer.20160763

Carrell, S. E., Hoekstra, M., & West, J. E. (2019). The Impact of College Diversity on Behavior toward Minorities. *American Economic Journal: Economic Policy*, *11*(4), 159–182. https://doi.org/10.1257/pol.20170069

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics*, *2*(1), 211–228. https://doi.org/10.1257/app.2.1.211

Cestau, D., Epple, D., & Sieg, H. (2017). Admitting Students to Selective Education Programs: Merit, Profiling, and Affirmative Action [Publisher: The University of Chicago Press]. *Journal of Political Economy*, *125*(3), 761–797. https://doi.org/10.1086/691702

Chetty, R., Deming, D. J., & Friedman, J. N. (2023). Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges. *NBER Working Paper*. https://doi.org/10.3386/w31492

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2020). Race and Economic Opportunity in the United States: An Intergenerational Perspective. *The Quarterly Journal of Economics*, *135*(2), 711–783. https://doi.org/10.1093/qje/qjz042

Corcoran, S. P., & Baker-Smith, E. C. (2018). Pathways to an Elite Education: Application, Admission, and Matriculation to New York City's Specialized High Schools. *Education Finance and Policy*, *13*(2), 256–279. https://doi.org/10.1162/edfp_a_00220

Cortes, K. E. (2010). Do bans on affirmative action hurt minority students? Evidence from the Texas Top 10% Plan. *Economics of Education Review*, *29*(6), 1110–1124. https://doi.org/10.1016/j.econedurev.2010.06.004

Cortes, K. E., & Klasik, D. (2020, December). *Uniform Admissions, Unequal Access: Did the Top 10% Plan Increase Access to Selective Flagship Institutions?* (Tech. rep. No. w28280). National Bureau of Economic Research. https://doi.org/10.3386/w28280

Crespin, R. (2022). *Essays on Public and Labor Economics* [Ph.D.]. Cornell University [ISBN: 9798351415314]. Retrieved June 26, 2025, from https://www.proquest.com/docview/2713545838/abstract/52F82DD6B61A44E4PQ/1

Cullen, J. B., Long, M. C., & Reback, R. (2013). Jockeying for position: Strategic high school choice under Texas' top ten percent plan. *Journal of Public Economics*, *97*, 32–48. https://doi.org/10.1016/j.jpubeco.2012.08.012

DCPS. (2023). *Selective High School Application Update for School Year 2022-2023* (tech. rep.). District of Columbia Public Schools. https://enrolldcps.dc.gov/sites/dcpsenrollment/files/page_content/attachments/English_SY22-23%20DCPS%20Selective%20HS%20Admissions%20Information%20Overview.pdf

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Dobbie, W., & Fryer, R. G. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, *6*(3), 58–75. https://doi.org/10.1257/app.6.3.58

East, C. N., Miller, S., Page, M., & Wherry, L. R. (2023). Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health. *American Economic Review*, *113*(1), 98–135. https://doi.org/10.1257/aer.20210937

Ellison, G., & Pathak, P. A. (2021). The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools. *American Economic Review*, *111*(3), 943–975. https://doi.org/10.1257/aer.20161290

Feerick Center for Social Justice. (2019). *Screened Out: The Lack of Access to NYC Screened Program Admissions Criteria* (tech. rep.). Fordham Law School. Retrieved August 27, 2024, from https://www.fordham.edu/media/home/schools/school-of-law/pdfs/screened-out-accessible.pdf

Finn, C. E., & Hockett, J. A. (2012, September). *Exam Schools*. Princeton University Press. Retrieved July 8, 2023, from https://press.princeton.edu/books/hardcover/9780691156675/exam-schools

Goodman, J., Gurantz, O., & Smith, J. (2020). Take Two! SAT Retaking and College Enrollment Gaps. *American Economic Journal: Economic Policy*, *12*(2), 115–158. https://doi.org/10.1257/pol.20170503

Goodman-Bacon, A. (2021). The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. *American Economic Review*, *111*(8), 2550–2593. https://doi.org/10.1257/aer.20171671

Hoxby, C. M. (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*.

Hoxby, C. M. (2025). Affirmative Action: Logic, Methods, and Evidence. Retrieved September 23, 2025, from https://conference.nber.org/conf_papers/f227376.slides.pdf

Hoxby, C. M., & Avery, C. (2013). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students. *Brookings Papers on Economic Activity*, *2013*, 1–65. Retrieved February 12, 2025, from https://www.brookings.edu/wp-content/uploads/2016/07/2013a_hoxby.pdf

Idoux, C. (2022). Can Selective School Admission Reforms Increase Student Achievement? *Working Paper*.

Kapor, A. (2020). Distributional Effects of Race-Blind Affirmative Action [Number: 2020-51 Publisher: Princeton University. Economics Department.]. *Working Papers*. Retrieved January 16, 2024, from https://ideas.repec.org//p/pri/econom/2020-51.html

Klasik, D., & Cortes, K. E. (2022). Uniform admissions, unequal access: Did the top 10% plan increase access to selective flagship institutions? *Economics of Education Review*, *87*, 102199. https://doi.org/10.1016/j.econedurev.2021.102199

Lawrence, E., & Mollborn, S. (2017). Racial/Ethnic Patterns of Kindergarten School Enrollment in the United States [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/socf.12352]. *Sociological Forum*, *32*(3), 635–658. https://doi.org/10.1111/socf.12352

Mader, N., Kramer, A., & Butel, A. (2018, November). *Promising Outcomes, Limited Potential: Diversity in Admissions in New York City Public Schools* (tech. rep.). Center for New York City Affairs. New York, NY.

Perez, J. (2018). Competition remains fierce for top programs as CPS sends out high school offers. *Chicago Tribune*. Retrieved August 27, 2024, from https://www.chicagotribune.com/2018/03/30/competition-remains-fierce-for-top-programs-as-cps-sends-out-high-school-offers/

Posey-Maddox, L., Powell, S. N., Roda, A., Lenhoff, S. W., & Miller, E. O. (2025). Advantaged Families' Opportunity Hoarding in U.S. K–12 Education: A Systematic Review of the Literature [Publisher: American Educational Research Association]. *Review of Educational Research*, 00346543241304766. https://doi.org/10.3102/00346543241304766

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, *109*(3), 774–809. https://doi.org/10.1257/aer.20180044

Rokkanen, M. (2015). *Extrapolation and bandwidth choice in the regression discontinuity design* [Doctoral dissertation, Massachusetts Institute of Technology]. Retrieved October 15, 2017, from https://dspace.mit.edu/handle/1721.1/90131

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2023.03.008

Rucinski, M., & Goodman, J. (2022). Racial Diversity and Measuring Merit: Evidence from Boston's Exam School Admissions. *Education Finance and Policy*, *17*(3), 408–431. https://doi.org/10.1162/edfp_a_00343

Sinclair, B., & Chen, C.-S. (2020, August). Understanding School Lunch Eligibility in the Common Core of Data. https://nces.ed.gov/blogs/nces/post/understanding-school-lunch-eligibility-in-the-common-core-of-data

Stuyvesant High School. (2023). Admissions. Retrieved January 22, 2024, from https://stuy.enschool.org/m/pages/index.jsp?uREC_ID=126615&type=d

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

U.S. Department of Education. (2023). Common Core of Data [Education Data Portal (Version 0.20.0), Urban Insitute Urban Institute, made available under the ODC Attribution License]. https://educationdata.urban.org/documentation/

U.S. News and World Report. (2023). Best U.S. High Schools (2022-2023). https://www.usnews.com/education/best-high-schools/national-rankings

Wing, C., Freedman, S. M., & Hollingsworth, A. (2024, January). Stacked Difference-in-Differences. https://doi.org/10.3386/w32054

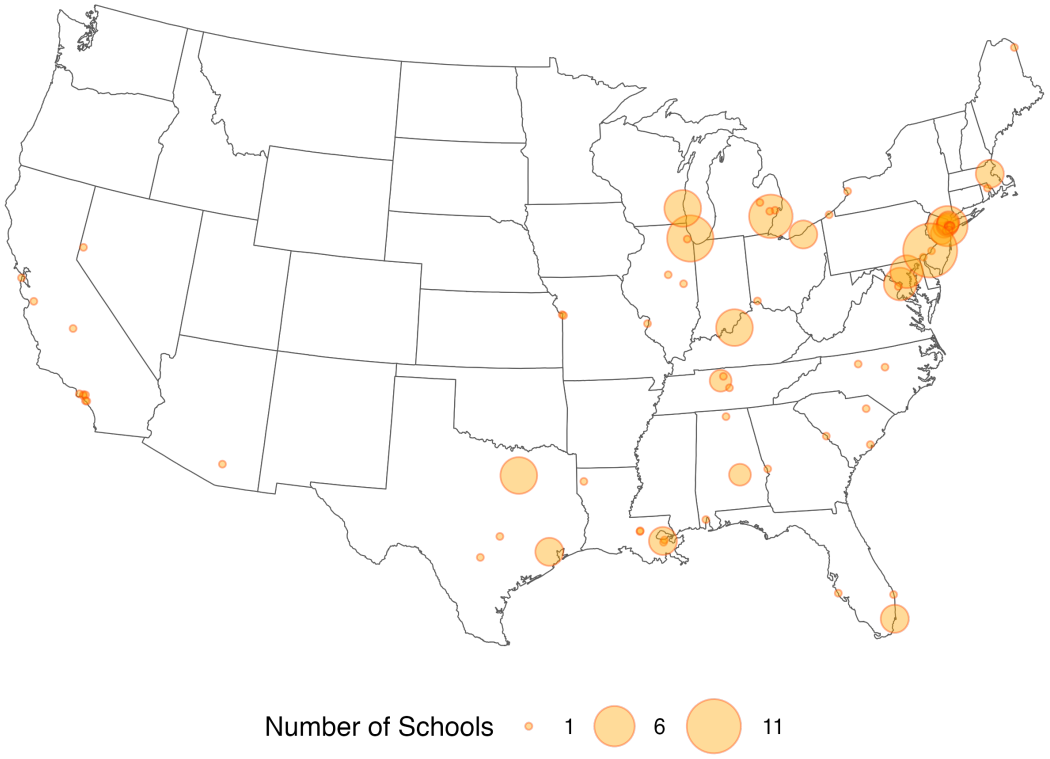
Woo, E. (1995). Caught on the Wrong Side of the Line?: Chinese Americans must outscore all other groups to enter elite Lowell High in San Francisco, sparking an ugly battle over diversity and the image of a 'model minority.' *Los Angeles Times*. Retrieved February 13, 2025, from https://www.proquest.com/docview/2021890024?sourcetype=Historical%20Newspapers

Woo, E. (1996). Racial Policy for Admission to S.F. School Altered [Num Pages: 2 Publisher: Los Angeles Times Communications LLC]. *Los Angeles Times*, NaN–NaN. Retrieved February 13, 2025, from https://www.proquest.com/docview/2047718203/abstract/357B736FE59B44B8PQ/1

Wooldridge, J. M. (2021, August). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. https://doi.org/10.2139/ssrn.3906345

Figures

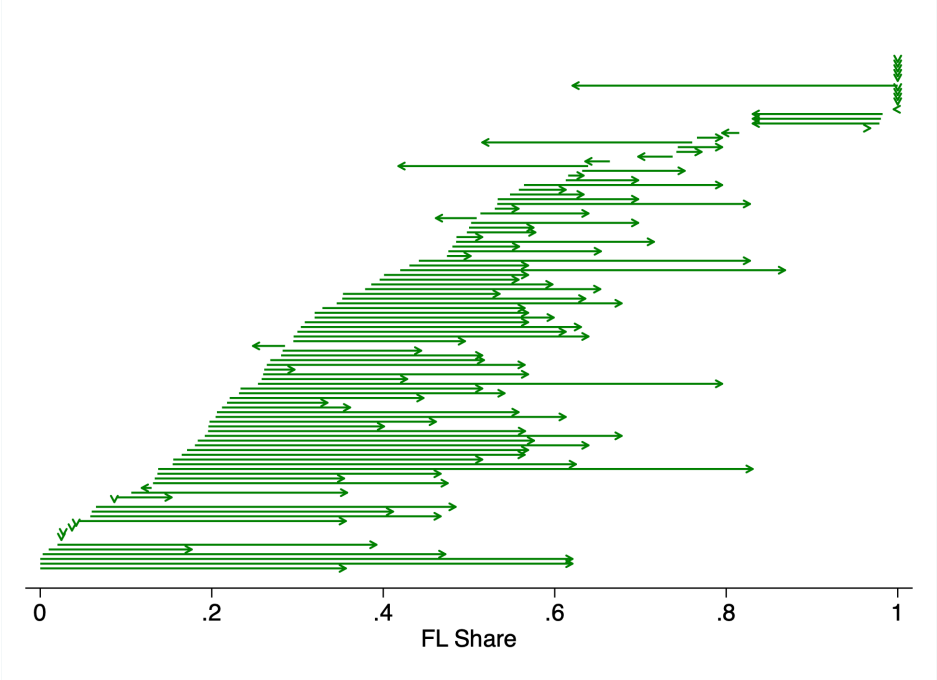
Figure 1: Map of school districts with at least one academically selective high school (ASHS)



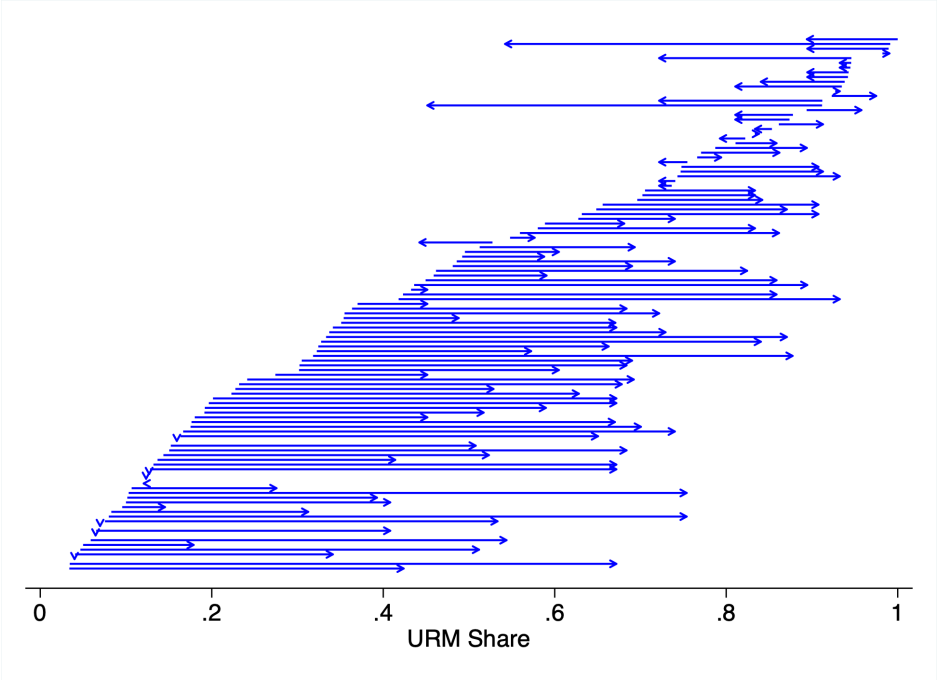
Notes: Each circle represents a separate school district. The circle size represents the number of 9th grade serving ASHSs in the corresponding school district between 2005 and 2021.

Figure 2: Gap Between ASHS and District Enrollment Share (typically underrepresented)

(a) FL Enrollment Share Gaps



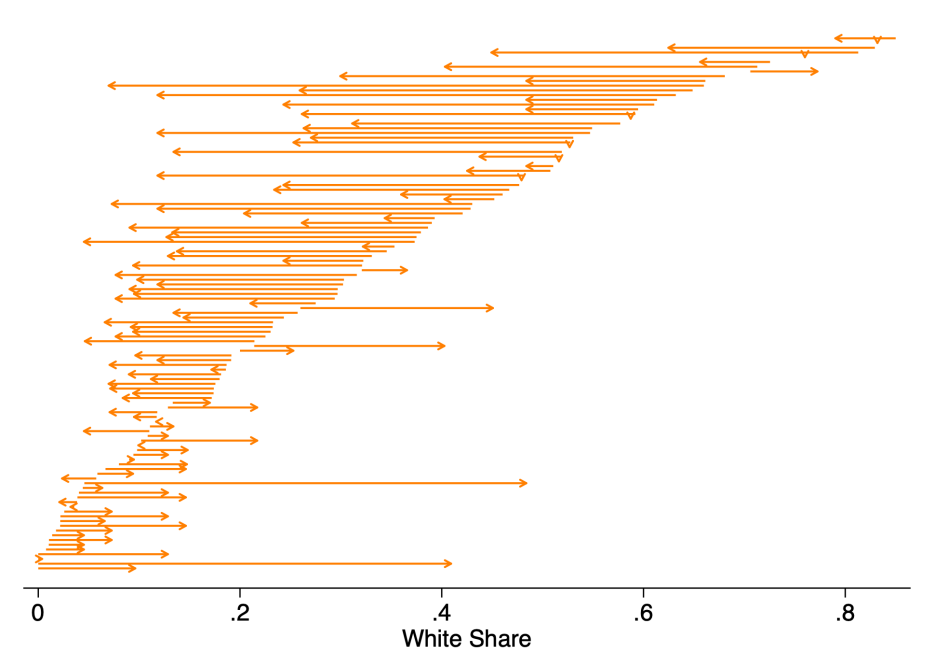
(b) URM Enrollment Share Gaps



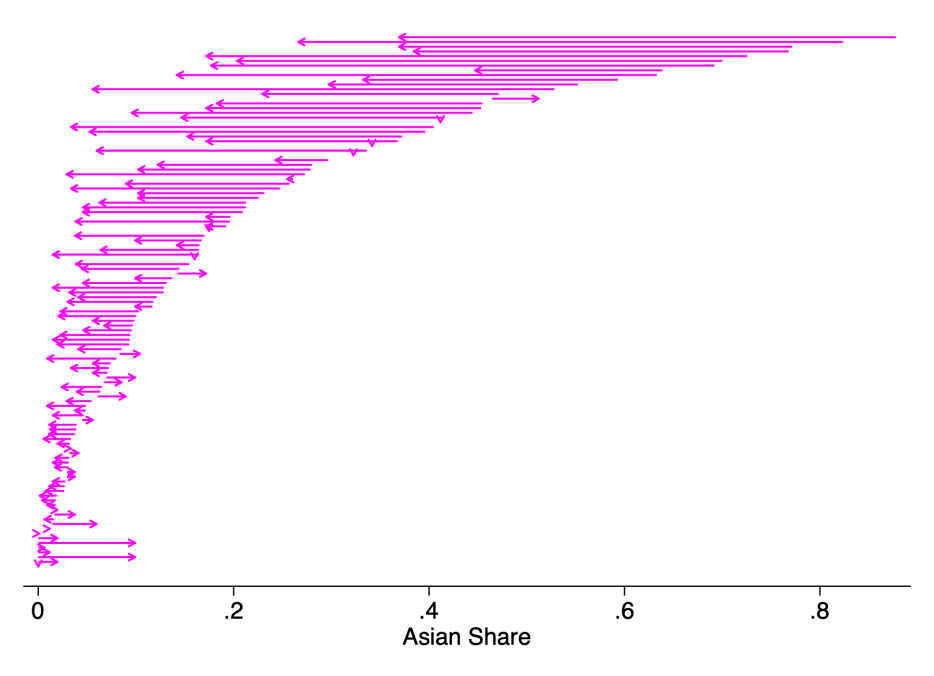
Notes: Based on sample of ASHSs in 2015-16. Each starting point represents each specific ASHS's actual enrollment shares for the corresponding group, while the arrow is the enrollment share for that same group in the corresponding school district.

Figure 3: Gap Between ASHS and District Enrollment Share (typically overrepresented)

(a) White Enrollment Share Gaps

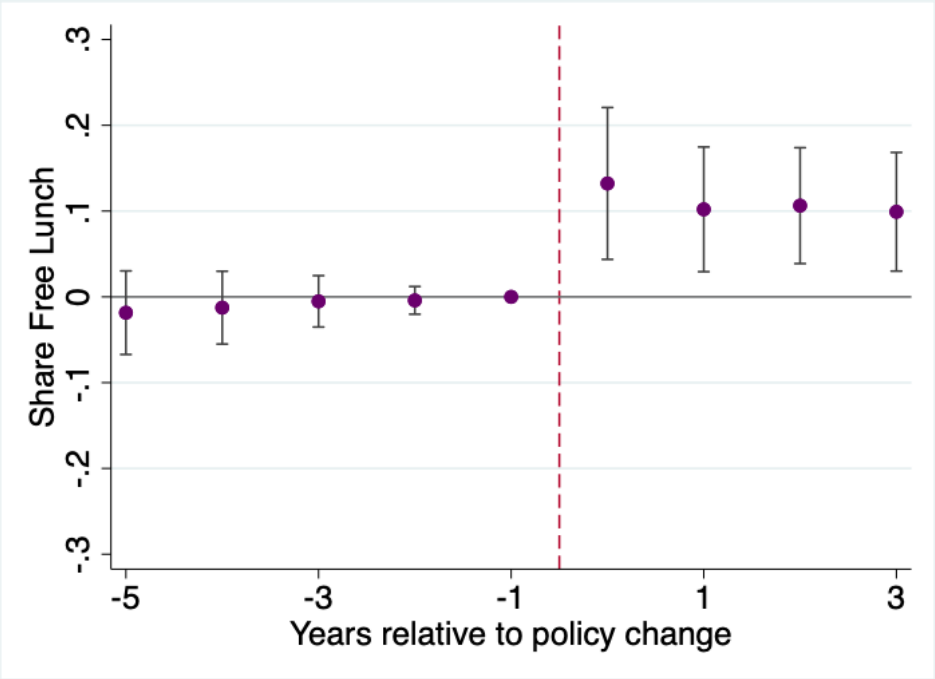


(b) Asian Enrollment Share Gaps



Notes: Based on sample of ASHSs in 2015-16. Each starting point represents each specific ASHS's actual enrollment shares for the corresponding group, while the arrow is the enrollment share for that same group in the corresponding school district.

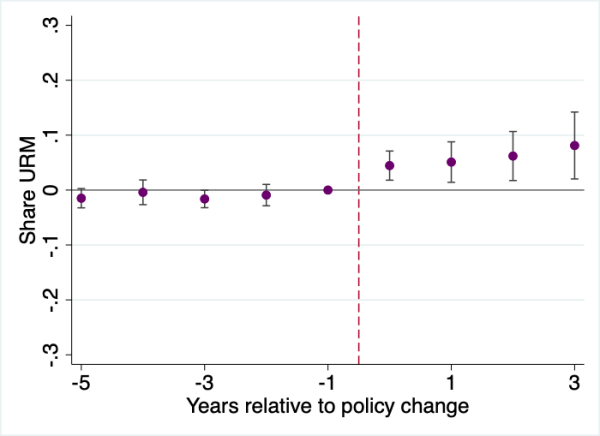
Figure 4: Effects on Free Lunch Enrollment Shares



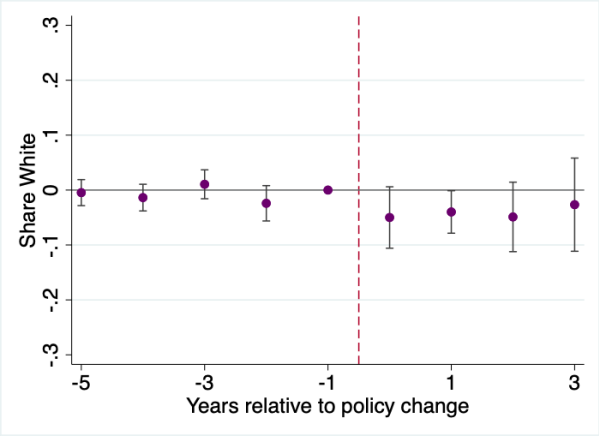
Notes: Event study is based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure 5: Effects on Racial/Ethnic Composition

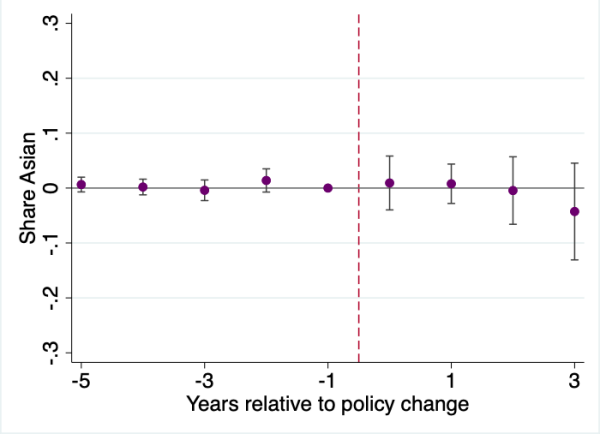
(a) URM Enrollment Shares



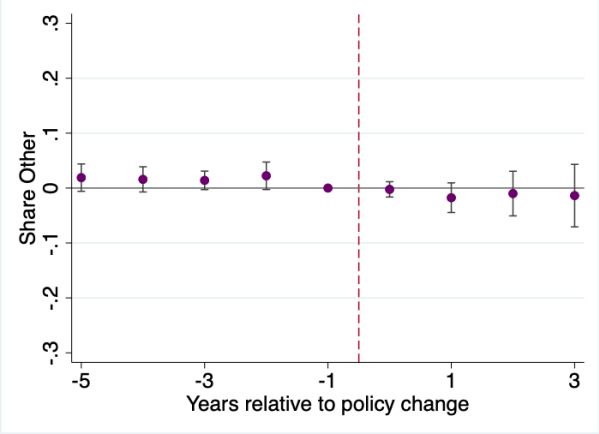
(b) White Enrollment Shares



(c) Asian Enrollment Shares

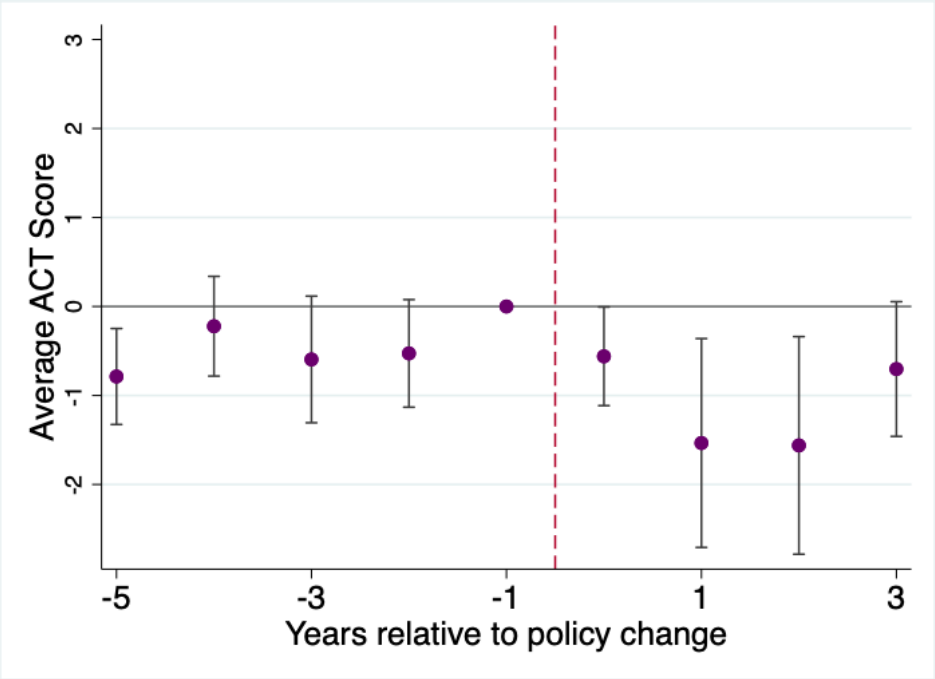


(d) Other Race Enrollment Shares



Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure 6: Effects on Average Composite ACT Score



Notes: Event study is based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. We bin all observations with relative time ≥ 4 for estimation, but we do not report the estimates due to lack of balance.

Tables

Table 1: Summary statistics of ASHSs (2015–16 SY)

	Full Sample	Final Analysis Sample		
	ASHSs	ASHSs	Control	Treated
	(1)	(2)	(3)	(4)
Panel A: Sample with available demographic data				
Free Lunch Shares	0.45	0.36	0.38	0.32
	(0.32)	(0.26)	(0.29)	(0.16)
URM Shares	0.48	0.42	0.51	0.23
	(0.32)	(0.30)	(0.31)	(0.19)
White Shares	0.30	0.32	0.30	0.36
	(0.23)	(0.23)	(0.24)	(0.20)
Asian Shares	0.19	0.22	0.16	0.36
	(0.21)	(0.23)	(0.21)	(0.21)
Other Race Shares	0.04	0.04	0.03	0.05
	(0.03)	(0.03)	(0.03)	(0.03)
9th Grade Enrollment	243.45	247.43	222.84	301.00
	(210.65)	(220.71)	(165.92)	(305.38)
School Enrollment	1025.72	1052.99	929.36	1322.32
	(811.34)	(847.67)	(619.15)	(1174.86)
N	130	89	61	28
Panel B: Subsample with available ACT/SAT data				
ACT Score	24.87	25.76	24.35	31.05
	(5.29)	(5.36)	(5.06)	(2.24)
N	117	63	50	13

Note: Summary statistics are based on sample described in Section 2.5. ACT score includes converted SAT scores as described in Section 2.5. Standard deviations are shown in parentheses.

Table 2: Effects of SES-based Admissions Policy on Demographic Enrollment Shares

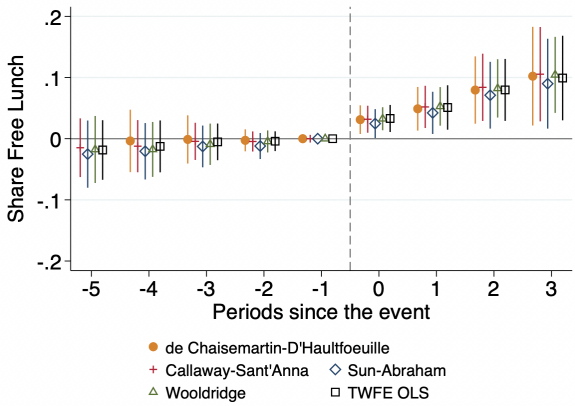
	Share Free Lunch	Share URM	Share White	Share Asian	Share Other	Avg. ACT Score
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall Sample						
SES Policy × Post	0.135**	0.065***	-0.042	-0.001	-0.024	
	(0.056)	(0.017)	(0.026)	(0.022)	(0.018)	
Percent Effect	39.8%	27.9%	-12.3%	-0.3%	-36.4%	
Observations	898	898	898	898	898	
Panel B: Test Scores Sample						
SES Policy × Post	0.141***	0.056***	-0.044*	0.001	-0.017	-0.529
	(0.052)	(0.014)	(0.023)	(0.017)	(0.015)	(0.326)
Percent Effect	41.7%	35.2%	-12.5%	0.2%	-27.0%	-1.7%
Observations	501	501	501	501	501	501

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

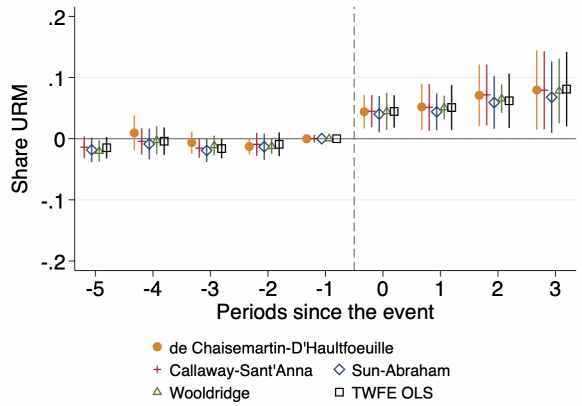
Appendix A Figures and Tables

Figure A1: Event Study Estimates of Impacts on School Demographics – Using various estimators (without dose-response)

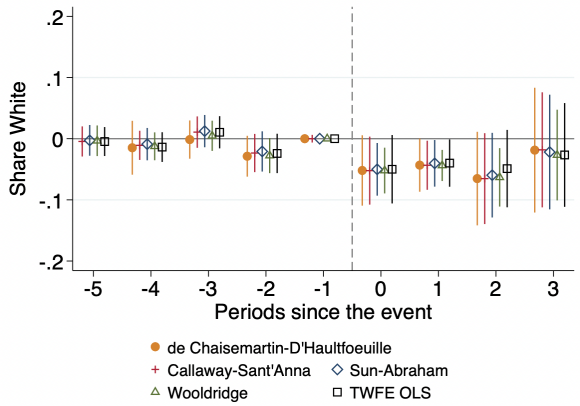
(a) Free Lunch Enrollment Shares



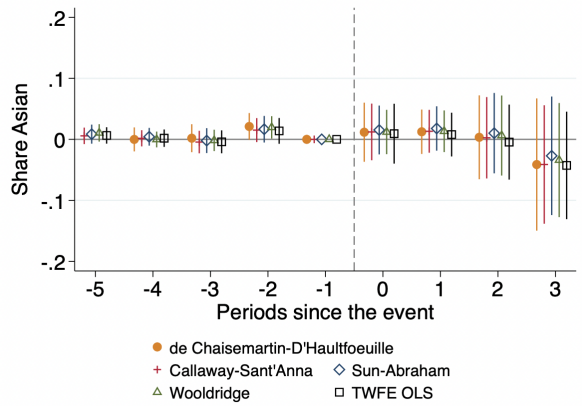
(b) URM Enrollment Shares



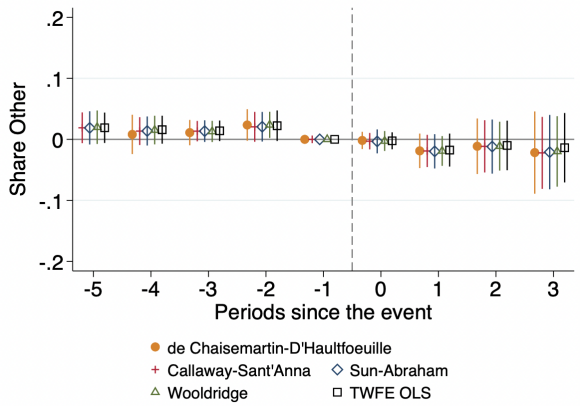
(c) White Enrollment Shares



(d) Asian Enrollment Shares



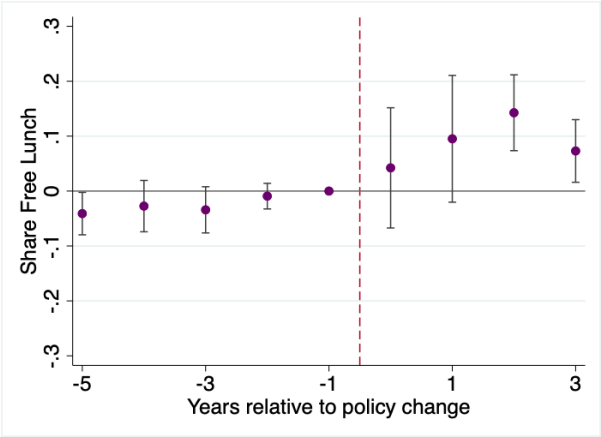
(e) Other Race Enrollment Shares



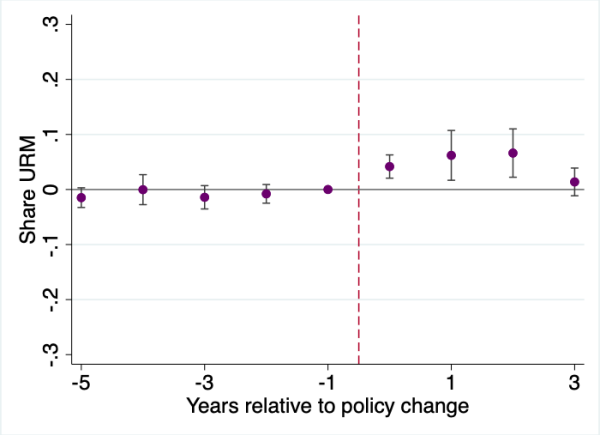
Notes: Each color/shape event study is based on a separate regression estimation (without dose response): black squares present TWFE estimates based on equation 2, orange circles present robust estimates by de Chaisemartin and D'Haultfœuille (2020), red crosses present robust estimates by Callaway and Sant'Anna (2021), blue diamonds present robust estimates by Sun and Abraham (2021), and green triangles present robust estimates by Wooldridge (2021). We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A2: Effects on Racial/Ethnic Composition— Test Scores Sample

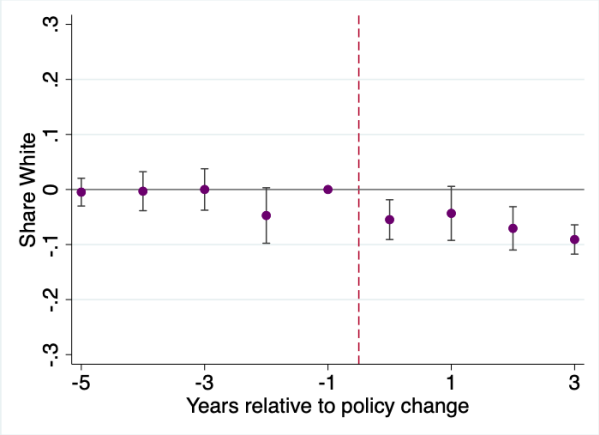
(a) Free Lunch Enrollment Shares



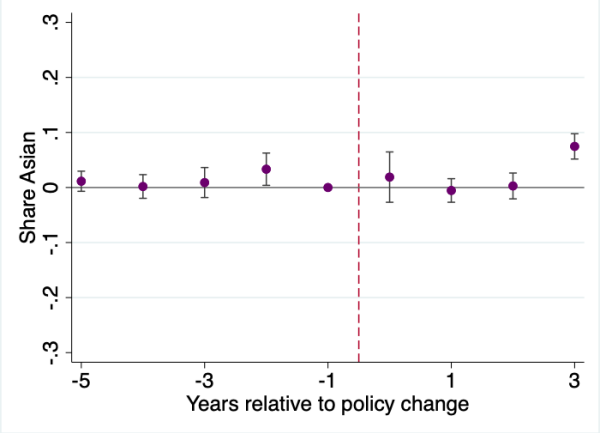
(b) URM Enrollment Shares



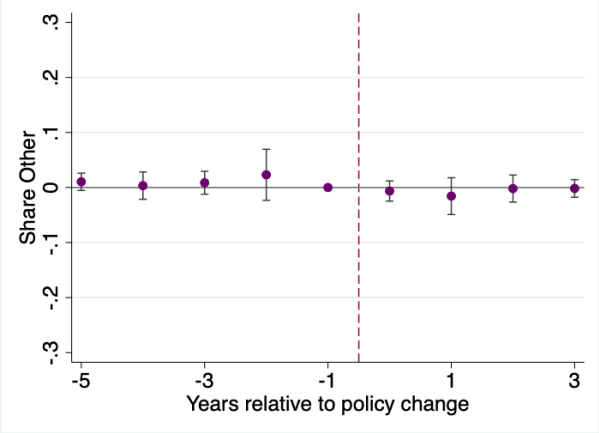
(c) White Enrollment Shares



(d) Asian Enrollment Shares

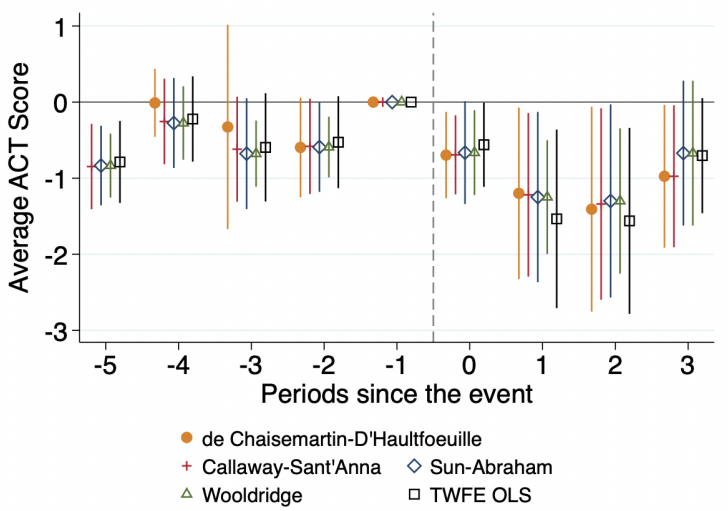


(e) Other Race Enrollment Shares



Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

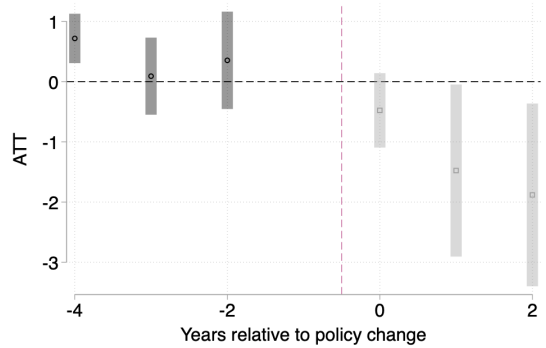
Figure A3: Event Study Estimates of Impacts on ACT Scores – Using various estimators



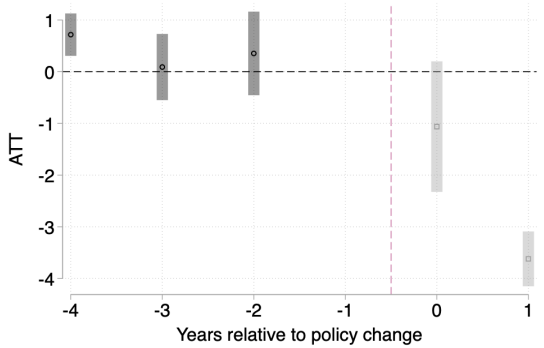
Notes: Each color/shape event study is based on a separate regression estimation (without dose response): black squares present TWFE estimates based on equation 2, orange circles present robust estimates by de Chaisemartin and D'Haultfœuille (2020), red crosses present robust estimates by Callaway and Sant'Anna (2021), blue diamonds present robust estimates by Sun and Abraham (2021), and green triangles present robust estimates by Wooldridge (2021).

Figure A4: Impacts on average ACT/SAT scores (focusing on NYC sample)

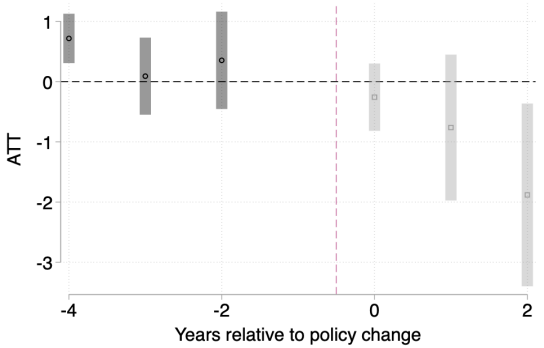
(a) Overall



(b) LIQ policies



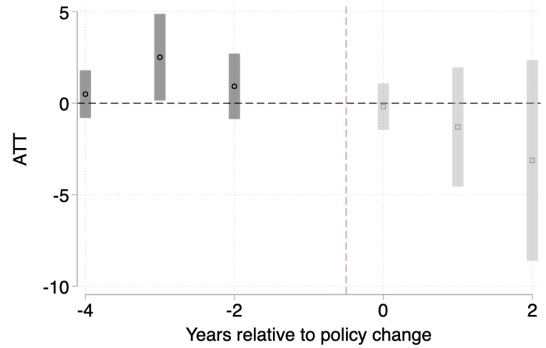
(c) SCA policies



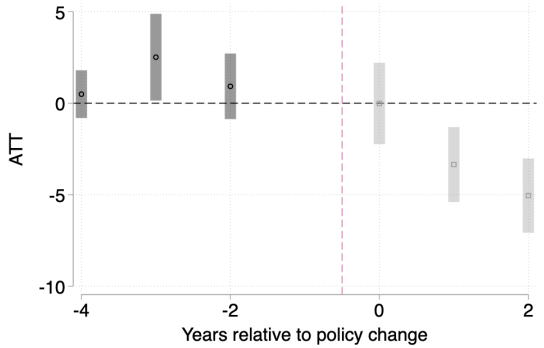
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3.

Figure A5: Impacts on average Algebra 1 regents exam score

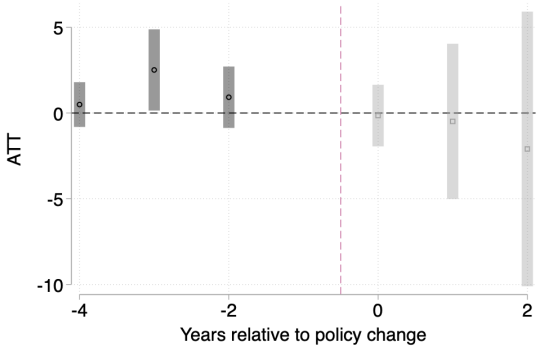
(a) Overall



(b) LIQ policies



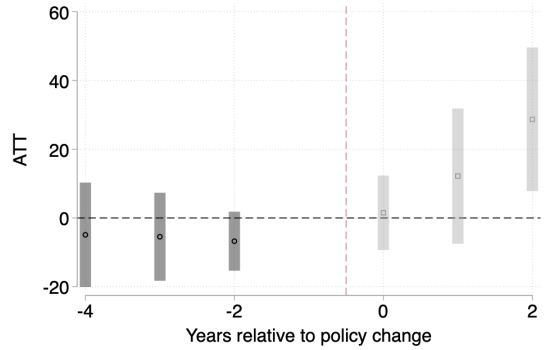
(c) SCA policies



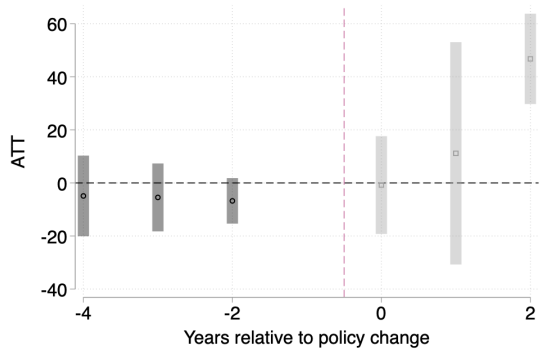
Notes: See notes in previous figure.

Figure A6: Impacts on Algebra 1 Participation Rates

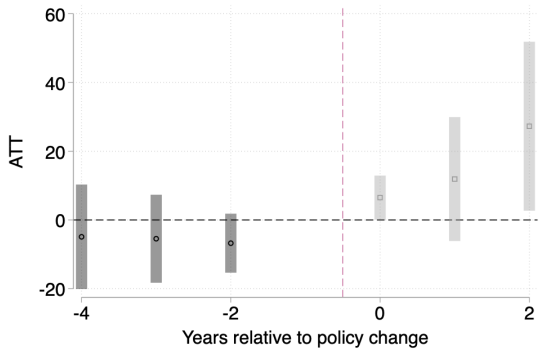
(a) Algebra 1 Participation Rate (Low-SES)



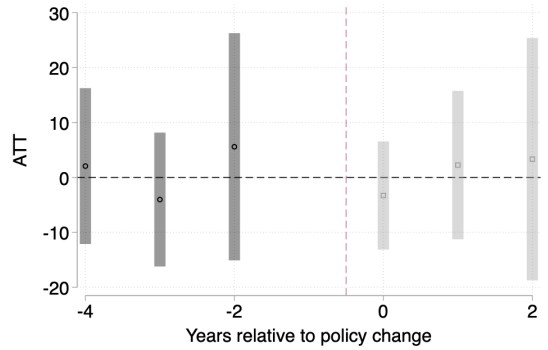
(b) Algebra 1 Participation Rate (Low-SES) (LIQ policies)



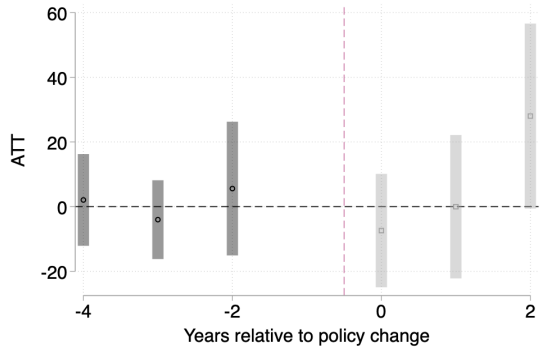
(c) Algebra 1 Participation Rate (Low-SES) (SCA policies)



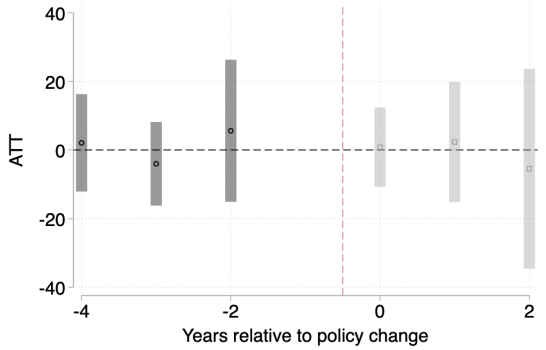
(d) Algebra 1 Participation Rate (non-low-SES)



(e) Algebra 1 Participation Rate (non-low-SES) (LIQ policies)



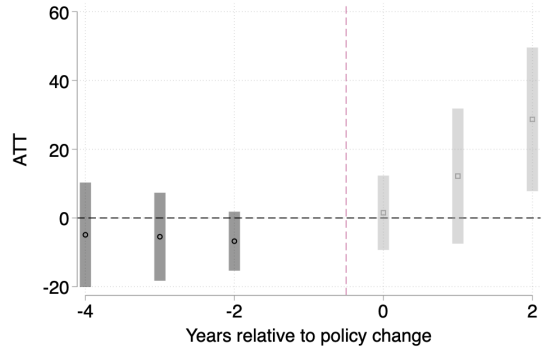
(f) Algebra 1 Participation Rate (non-low-SES) (SCA policies)



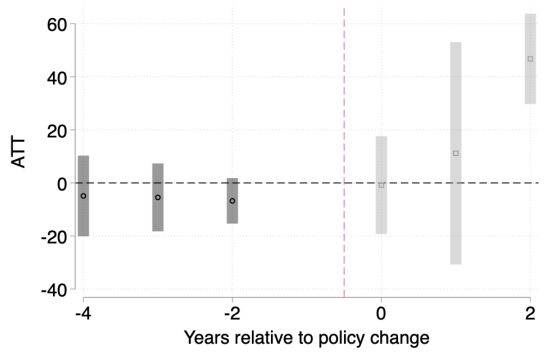
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A7: Impacts on Algebra 1 Participation Rates

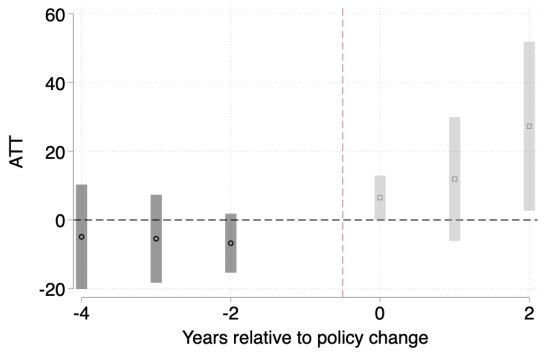
(a) Algebra 1 Participation Rate (Low-SES)



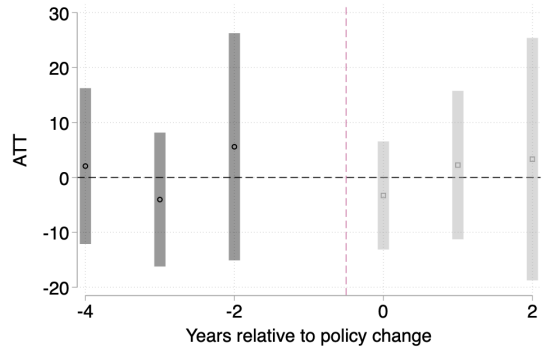
(b) Algebra 1 Participation Rate (Low-SES) (LIQ policies)



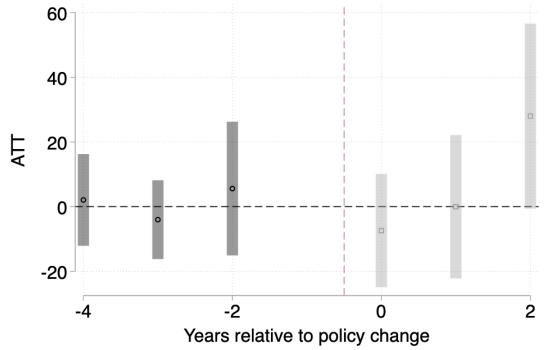
(c) Algebra 1 Participation Rate (Low-SES) (SCA policies)



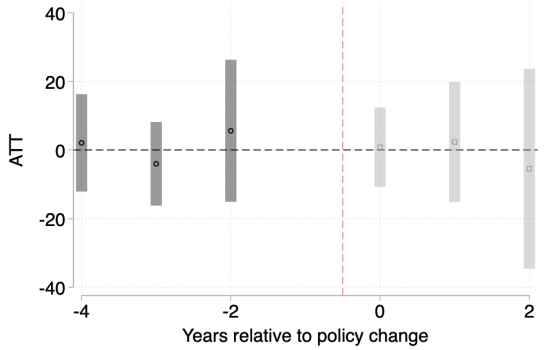
(d) Algebra 1 Participation Rate (non-low-SES)



(e) Algebra 1 Participation Rate (non-low-SES) (LIQ policies)



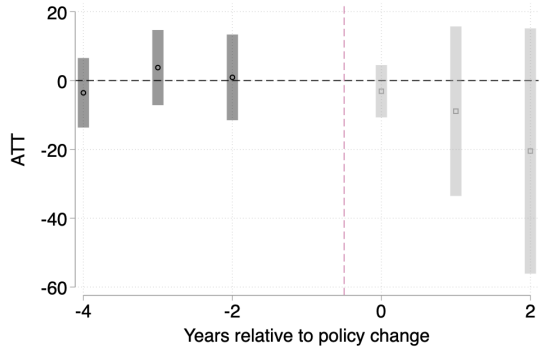
(f) Algebra 1 Participation Rate (non-low-SES) (SCA policies)



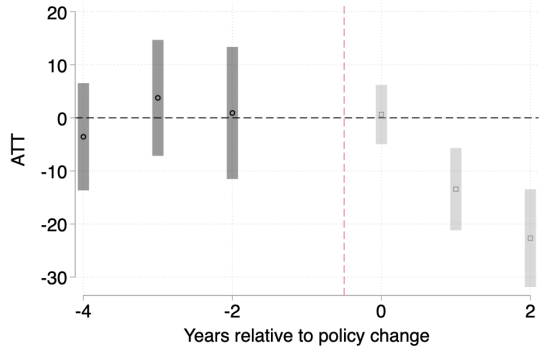
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A8: Algebra 1 Highest-Performance Rate

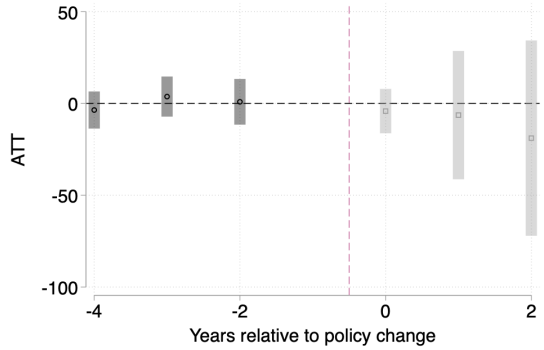
(a) Algebra 1 Highest-Performance Rate (Low-SES)



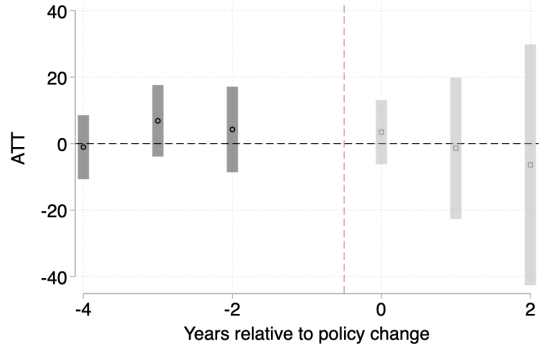
(b) Algebra 1 Highest-Performance Rate (Low-SES) (LIQ policies)



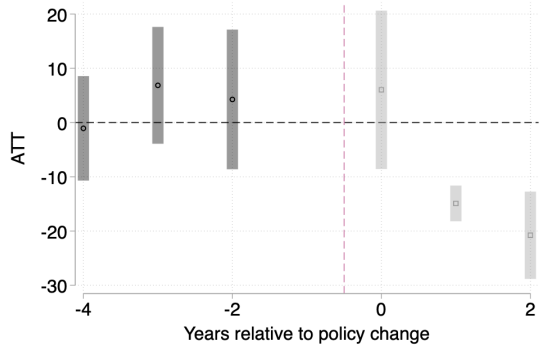
(c) Algebra 1 Highest-Performance Rate (Low-SES) (SCA policies)



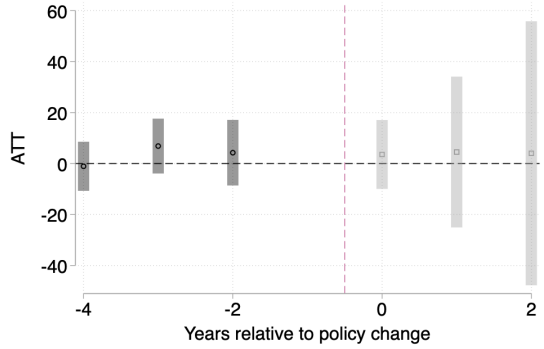
(d) Algebra 1 Highest-Performance Rate (non-low-SES)



(e) Algebra 1 Highest-Performance Rate (non-low-SES) (LIQ policies)



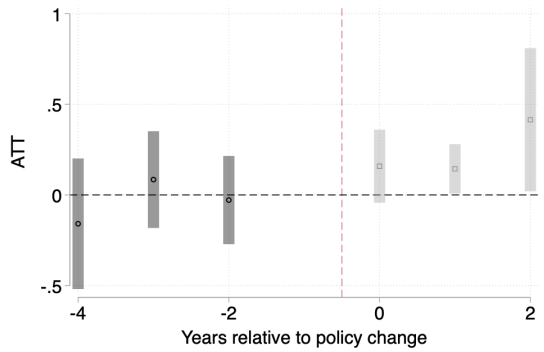
(f) Algebra 1 Highest-Performance Rate (non-low-SES) (SCA policies)



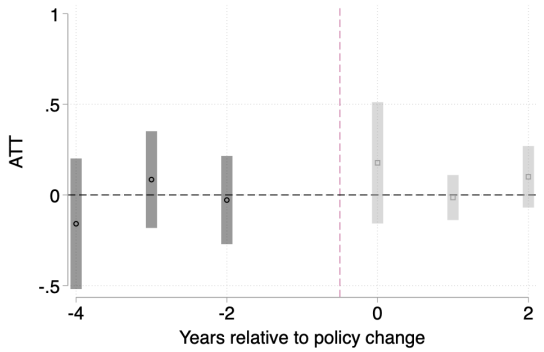
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A9: Suspensions and Race-based Harassment/Bullying Incidents

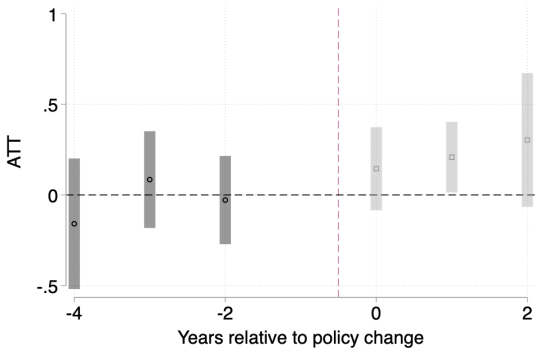
(a) % Students Suspended



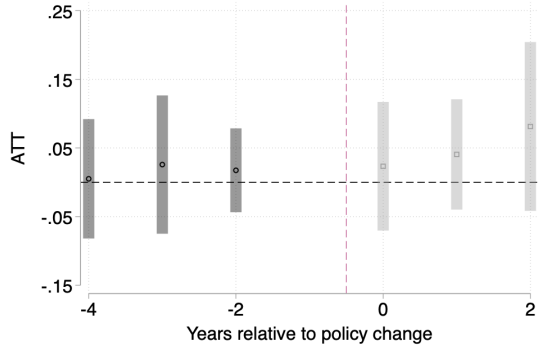
(b) % Students Suspended (LIQ policies)



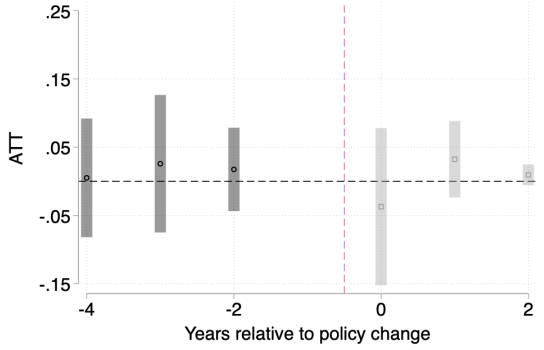
(c) % Students Suspended (SCA policies)



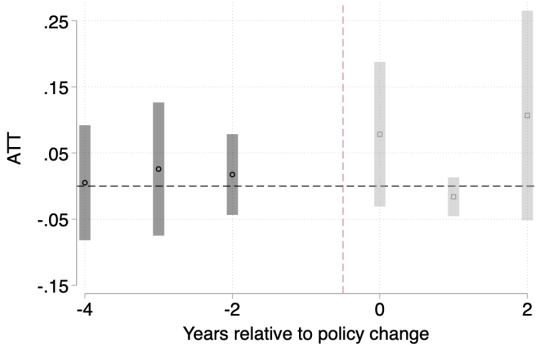
(d) Incidents of Racial Discrimination



(e) Incid. of Racial Disc. (LIQ policies)

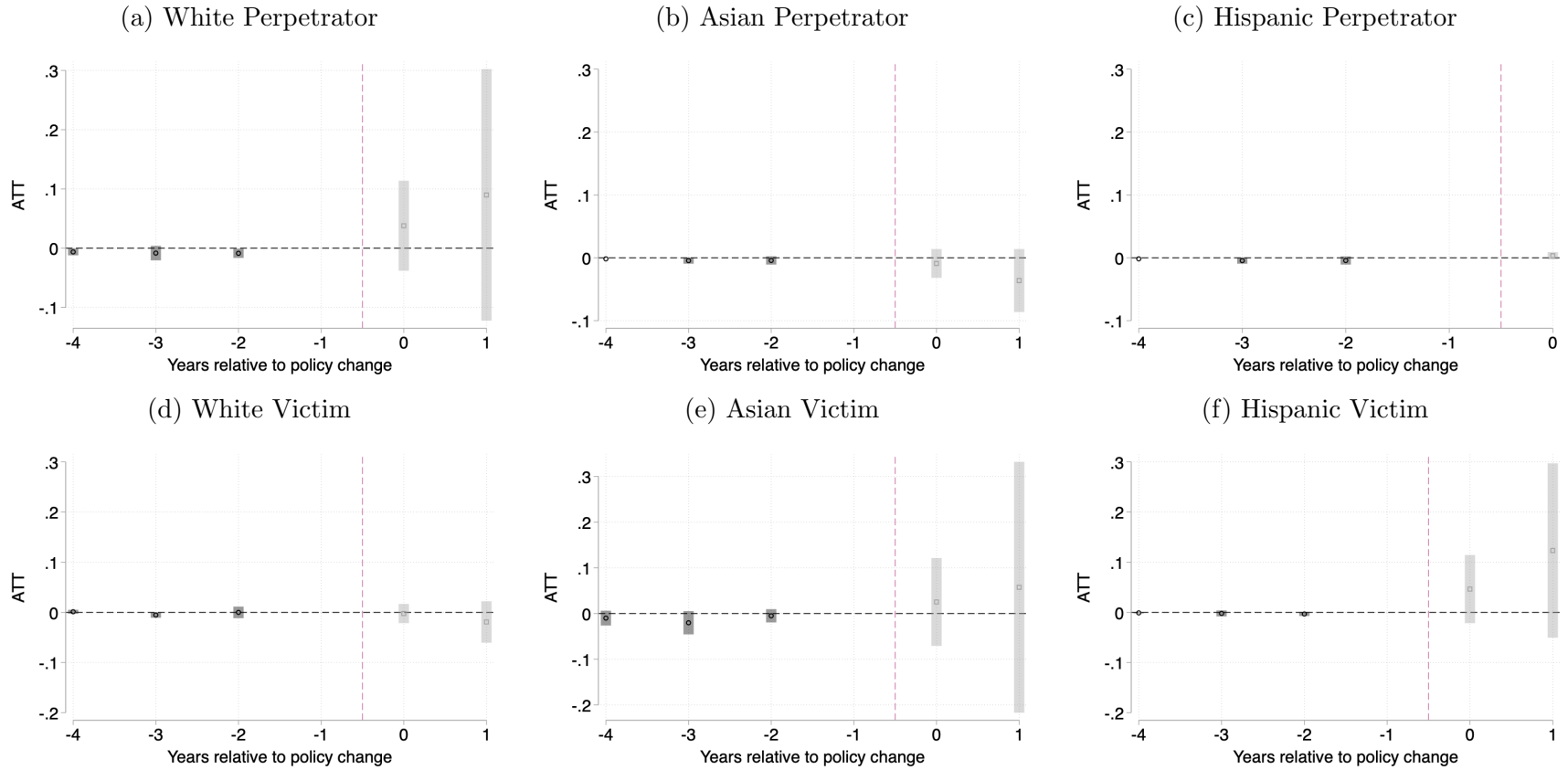


(f) Incid. of Racial Disc. (SCA policies)



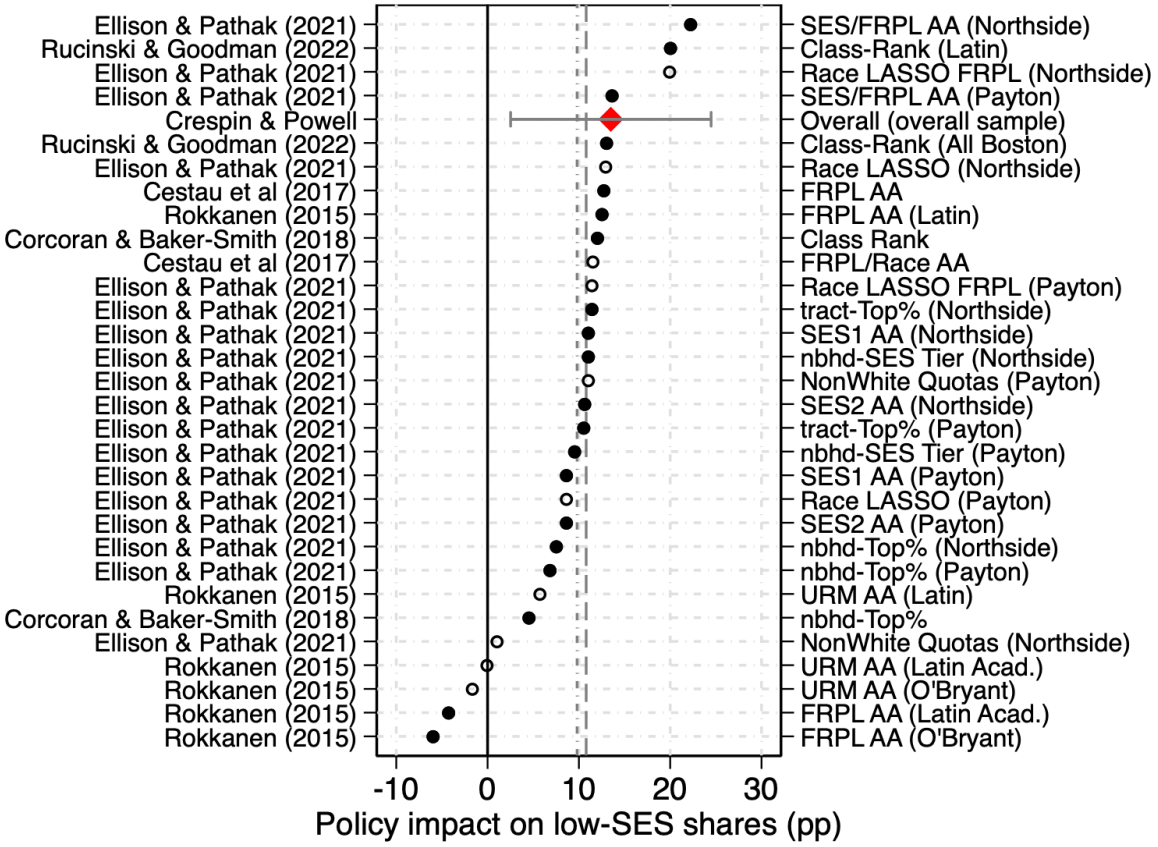
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A10: Incidents of Racial Discrimination, by race of perpetrator or victim



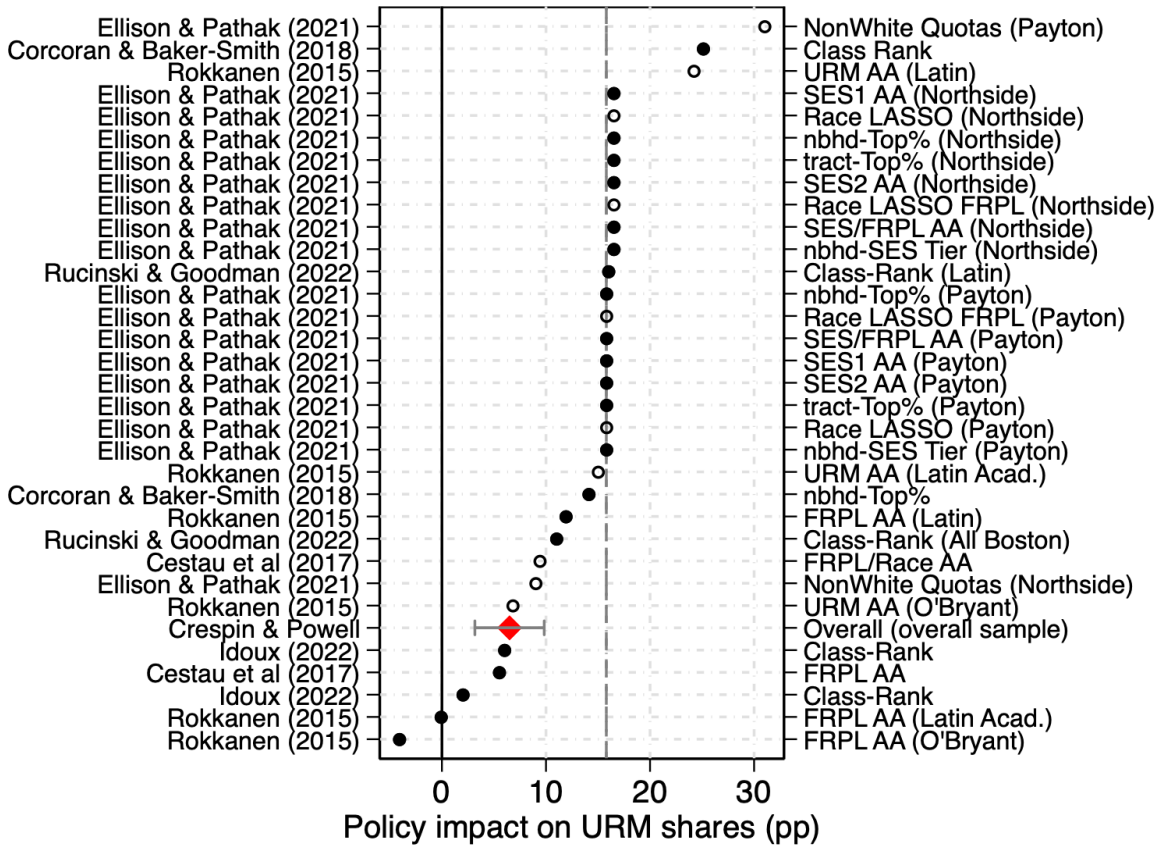
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A11: Forest Plot, Policy Impacts on low-SES Enrollment Shares



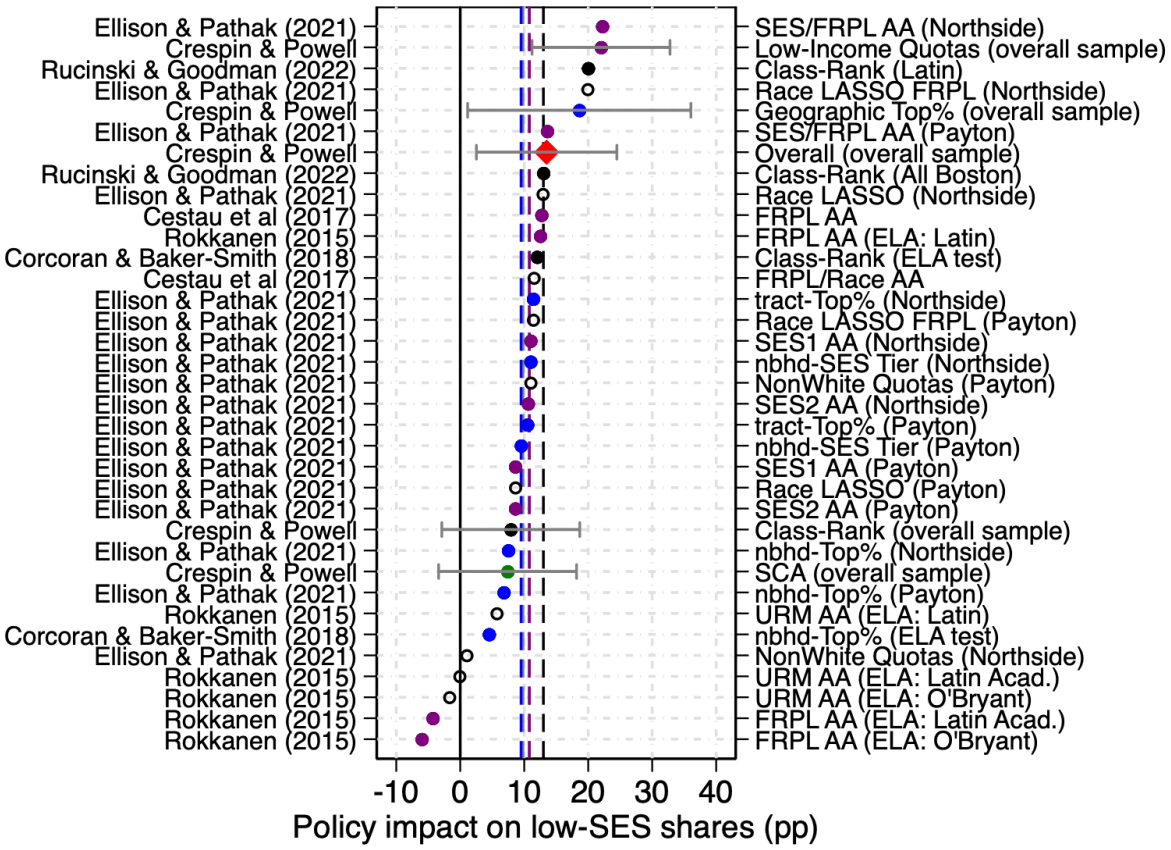
Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates of race-neutral AA policies, while the dotted gray line presents predictions from race-based AA policies.

Figure A12: Forest Plot, Policy Impacts on URM Enrollment Shares



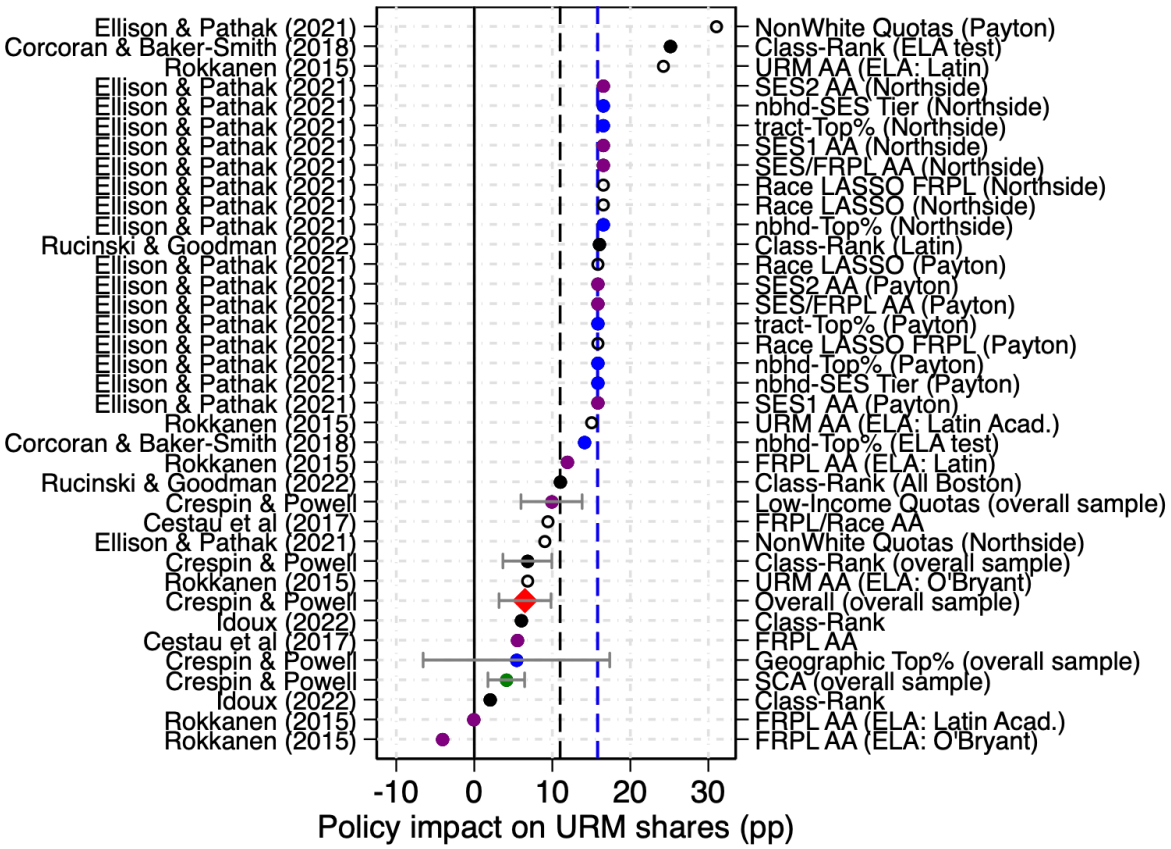
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates of race-neutral AA policies, while the dotted gray line presents predictions from race-based AA policies

Figure A13: Forest Plot, Policy Impacts on low-SES Enrollment Shares



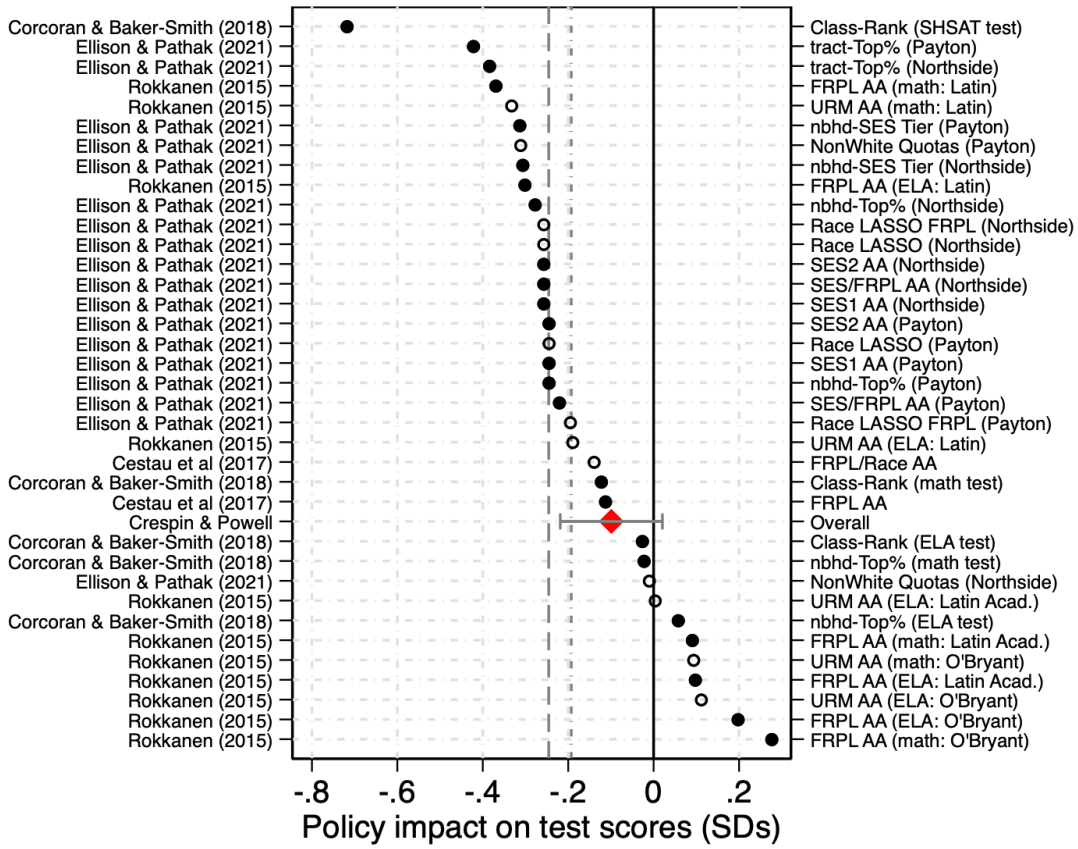
Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate.

Figure A14: Forest Plot, Policy Impacts on URM Enrollment Shares



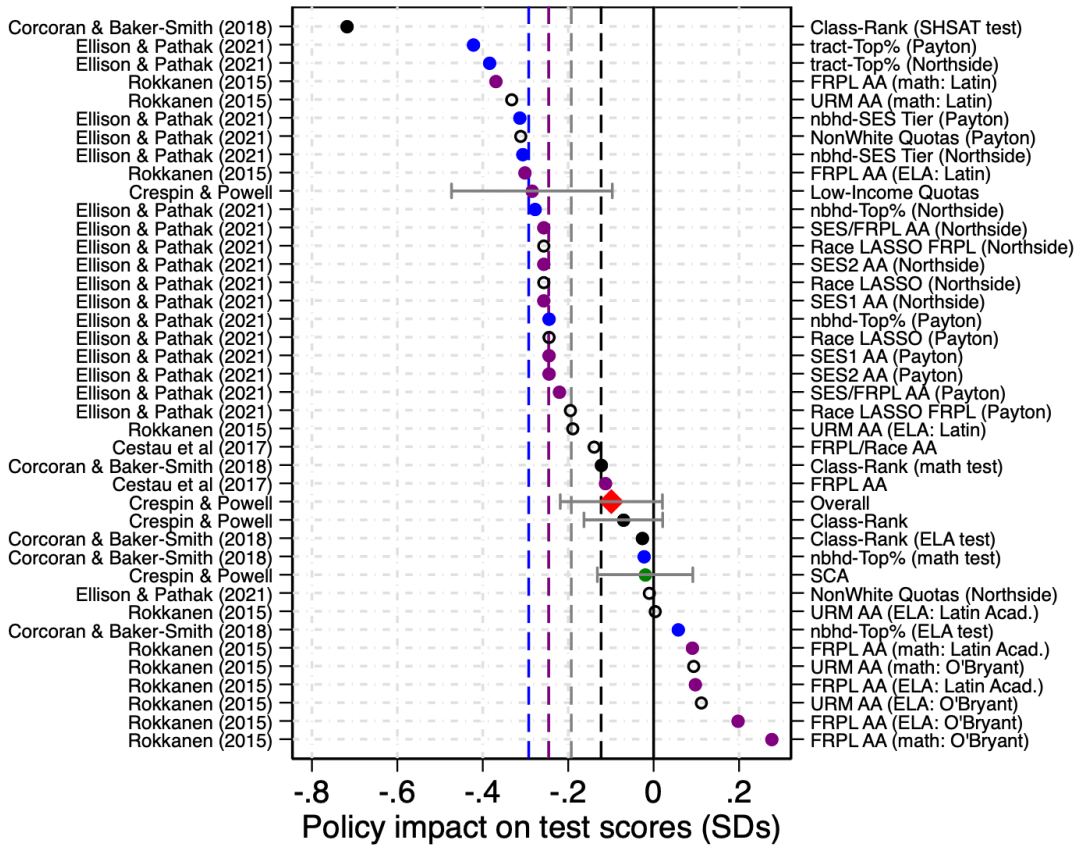
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate.

Figure A15: Forest Plot, Policy Impacts on Average Test Scores



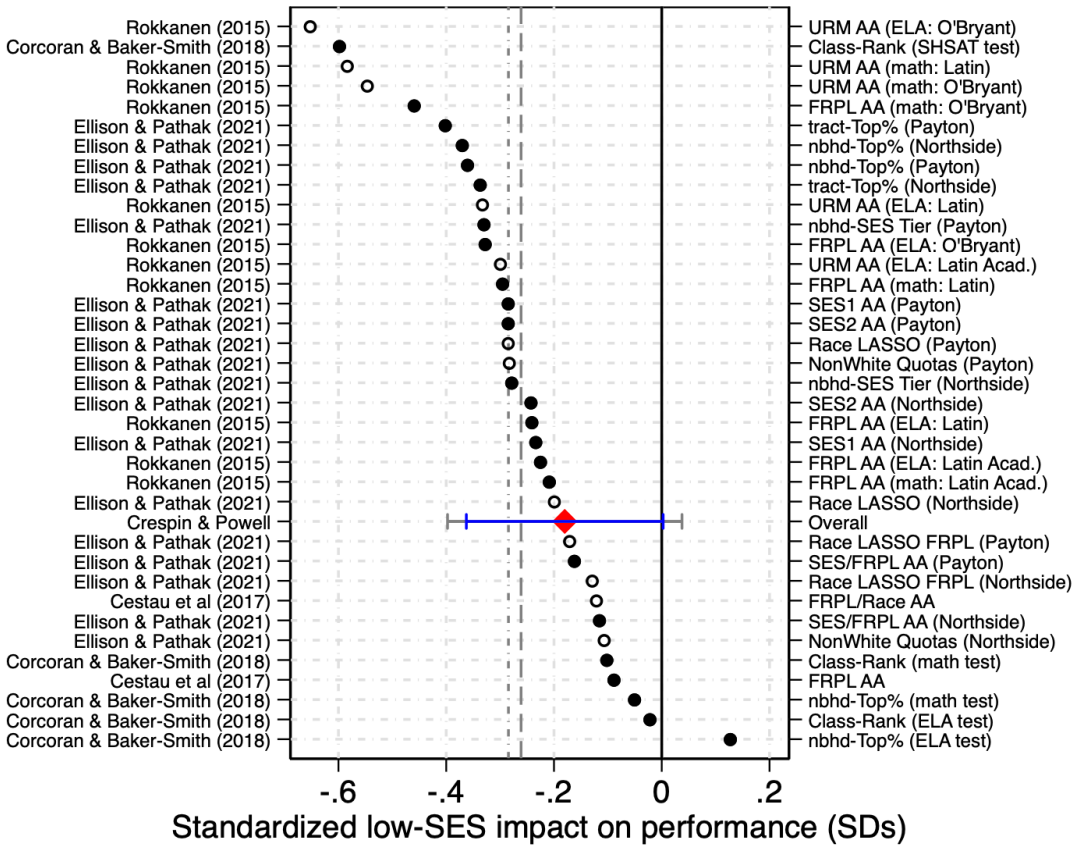
Notes: Each estimate represents the estimated policy impact on average test scores. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate. Filled points represent simulation-based race-neutral policy estimates, while hollow points are based on race-based policy simulations. The vertical dashed gray line represents the median for the simulation-based estimates of race-neutral AA policies, while the dotted gray line presents predictions from race-based AA policies.

Figure A16: Forest Plot, Policy Impacts on Average Test Scores (by specific policy type)



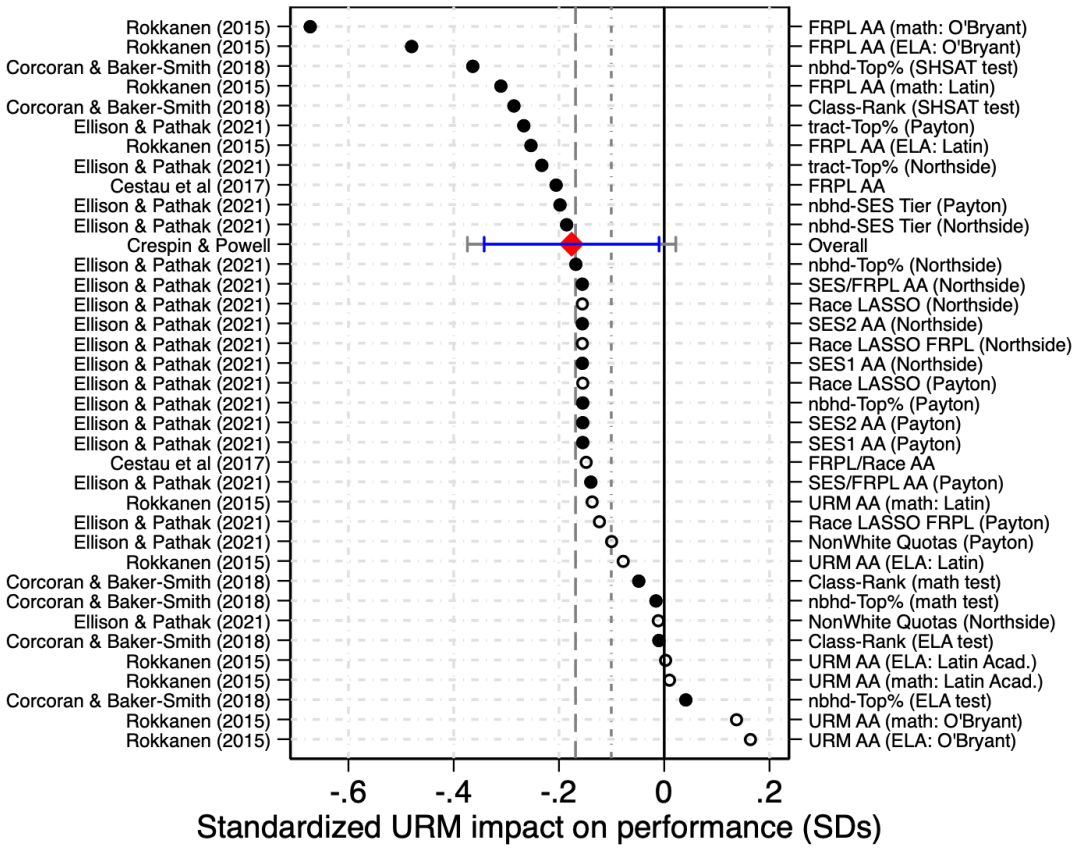
Notes: Each estimate represents the estimated policy impact on average test scores. The right-side y-axis categorizes the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate.

Figure A17: Forest Plot, FL Enrollment Impacts on Academic Achievement



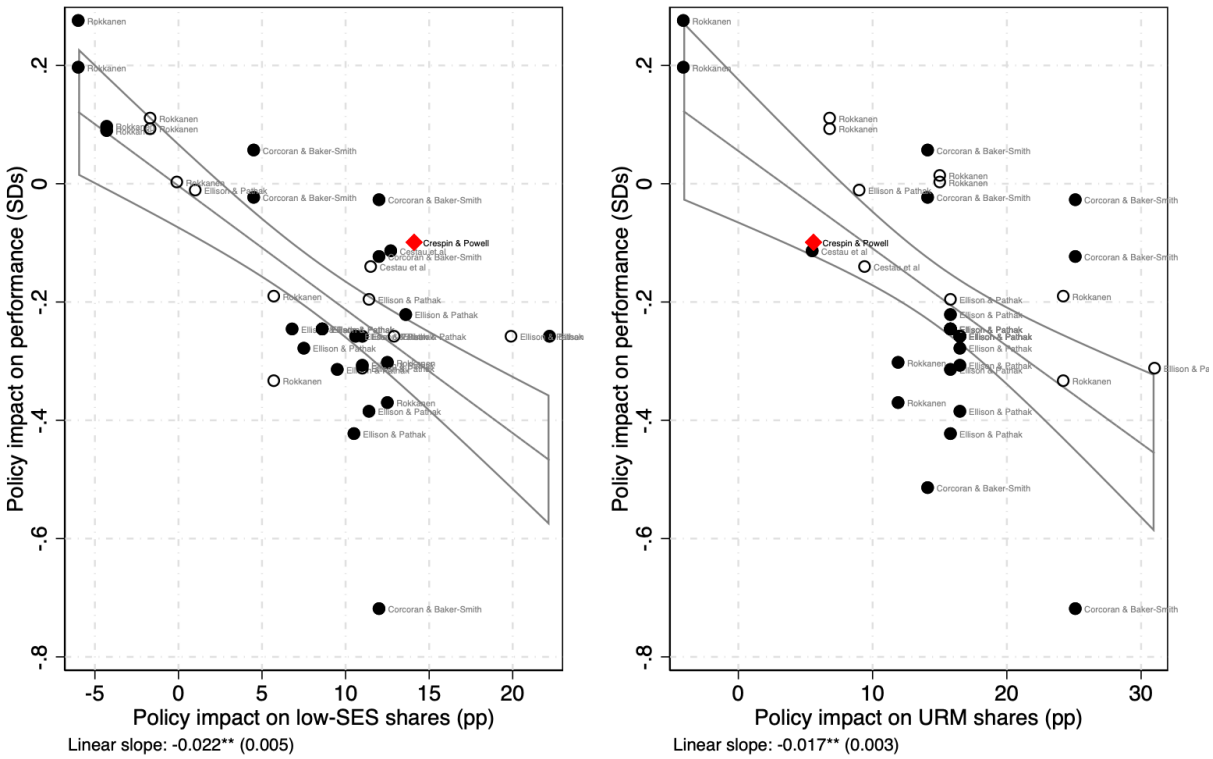
Notes: Each estimate represents the impact of a 10 percentage point increase in FL enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates of race-neutral AA policies, while the dotted gray line presents predictions from race-based AA policies.

Figure A18: Forest Plot, URM Enrollment Impacts on Academic Achievement



Notes: Each estimate represents the impact of a 10 percentage point increase in URM enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates of race-neutral AA policies, while the dotted gray line presents predictions from race-based AA policies.

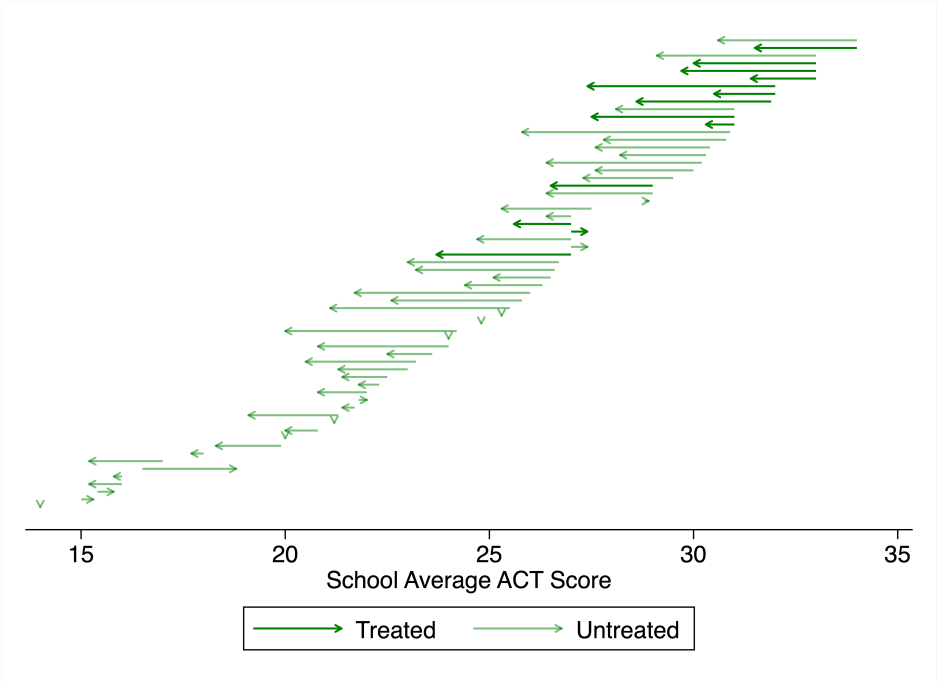
Figure A19: Comparing FL/Performance tradeoffs to URM/Performance tradeoffs



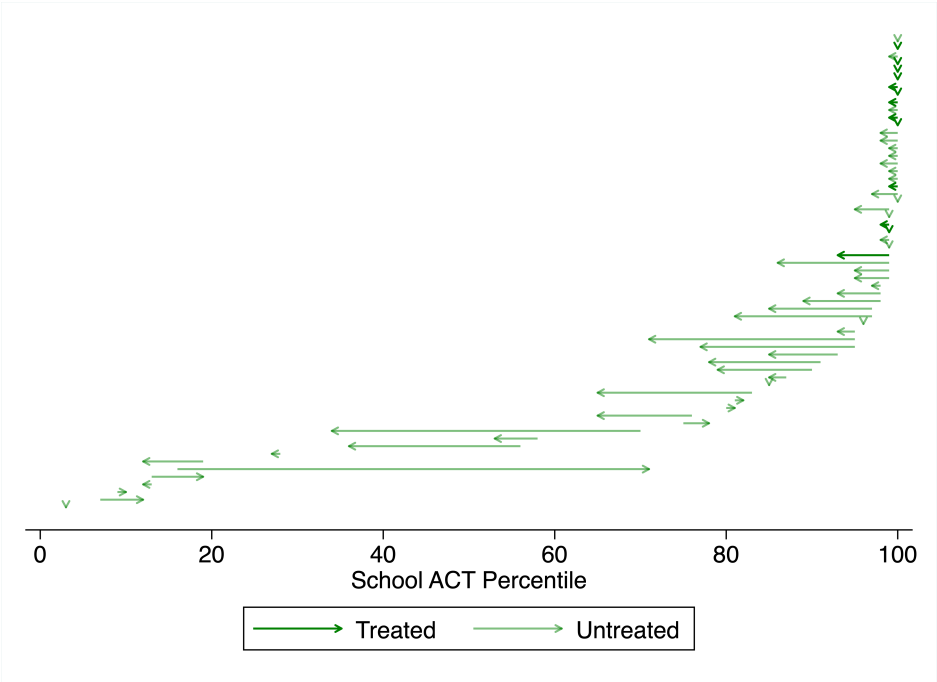
Notes: Each coordinate represents the policy impact on academic achievement (y-axis) and on low-SES or URM enrollment shares (x-axis) for each study-policy. Point labels represent authors. The red diamonds represent our own estimated policy impacts. Filled circles represent race-neutral AA policy simulation predictions, while hollow circles are race-based policies. The linear slopes are estimated using OLS regressions with standard errors clustered by study.

Figure A20: Simulated Effect of Closing FL Representation Gaps on Test Scores

(a) Test Score Effect



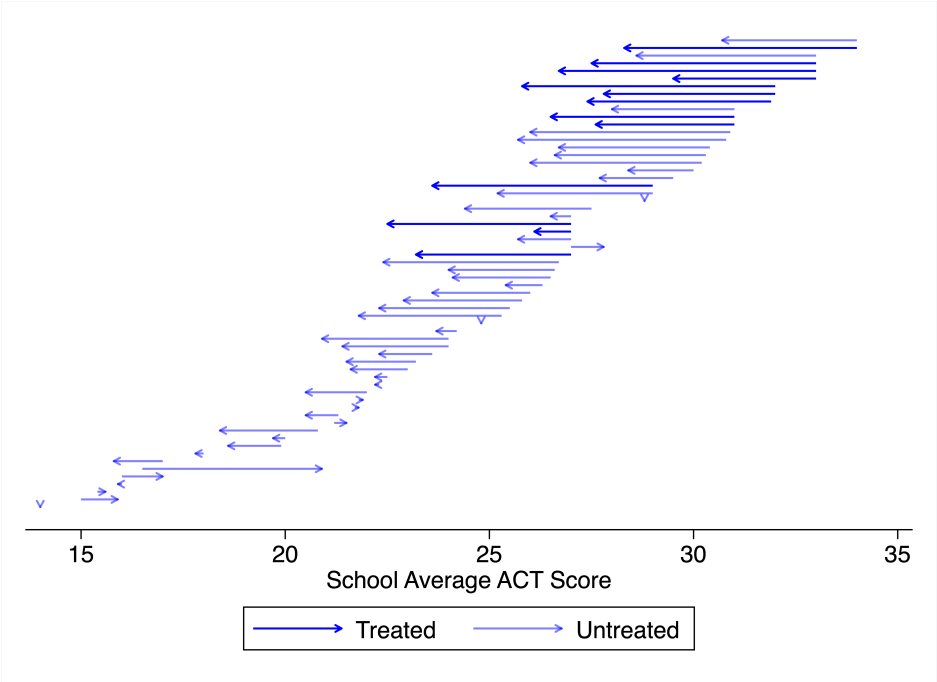
(b) Test Percentile Effect



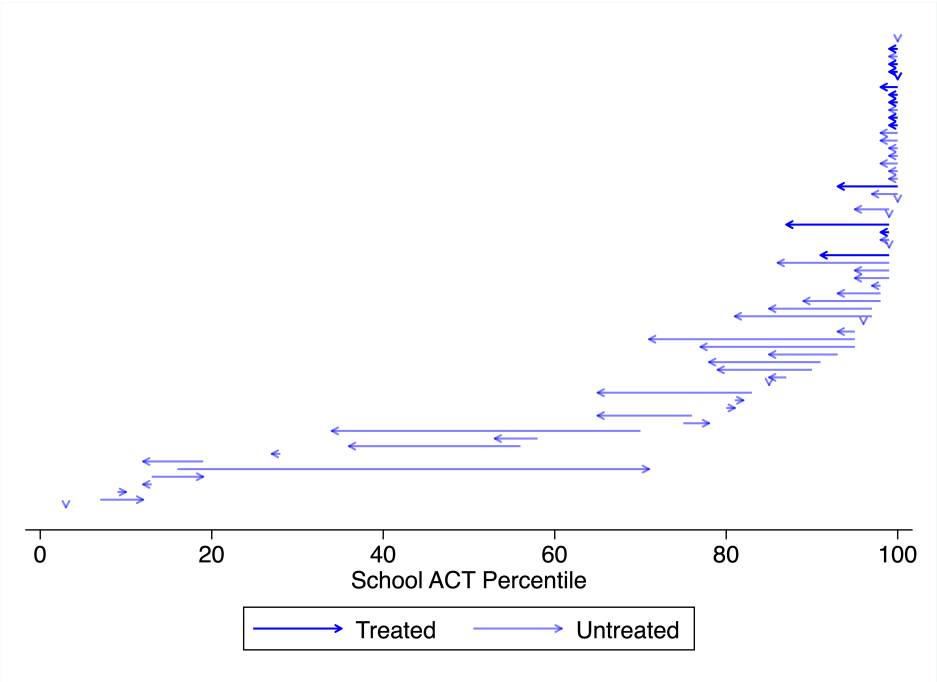
Notes: Based on sample of untreated ASHSs in 2015-16 and treated ASHSs in the year prior to treatment. Each starting point represents each specific ASHS's actual average ACT score for the corresponding group, while the arrow is the simulated average score after matching district free lunch share.

Figure A21: Simulated Effect of Closing URM Representation Gaps on Test Scores

(a) Test Score Effect

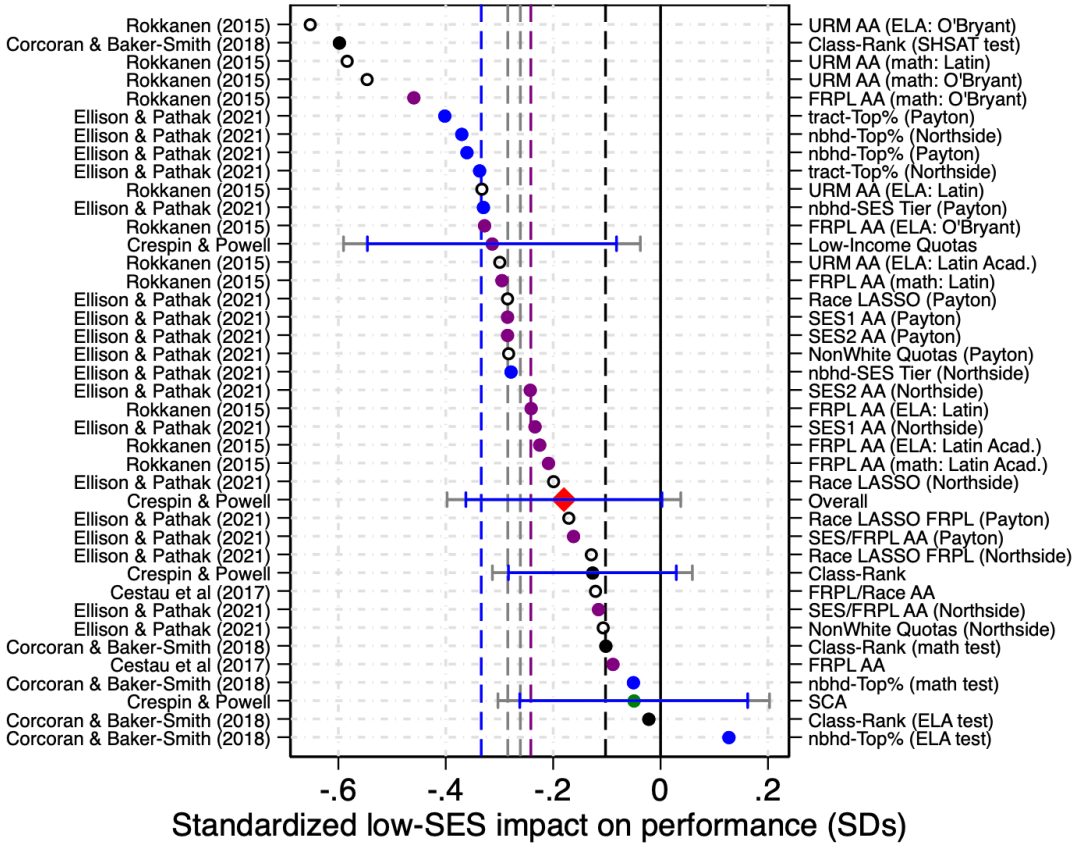


(b) Test Percentile Effect



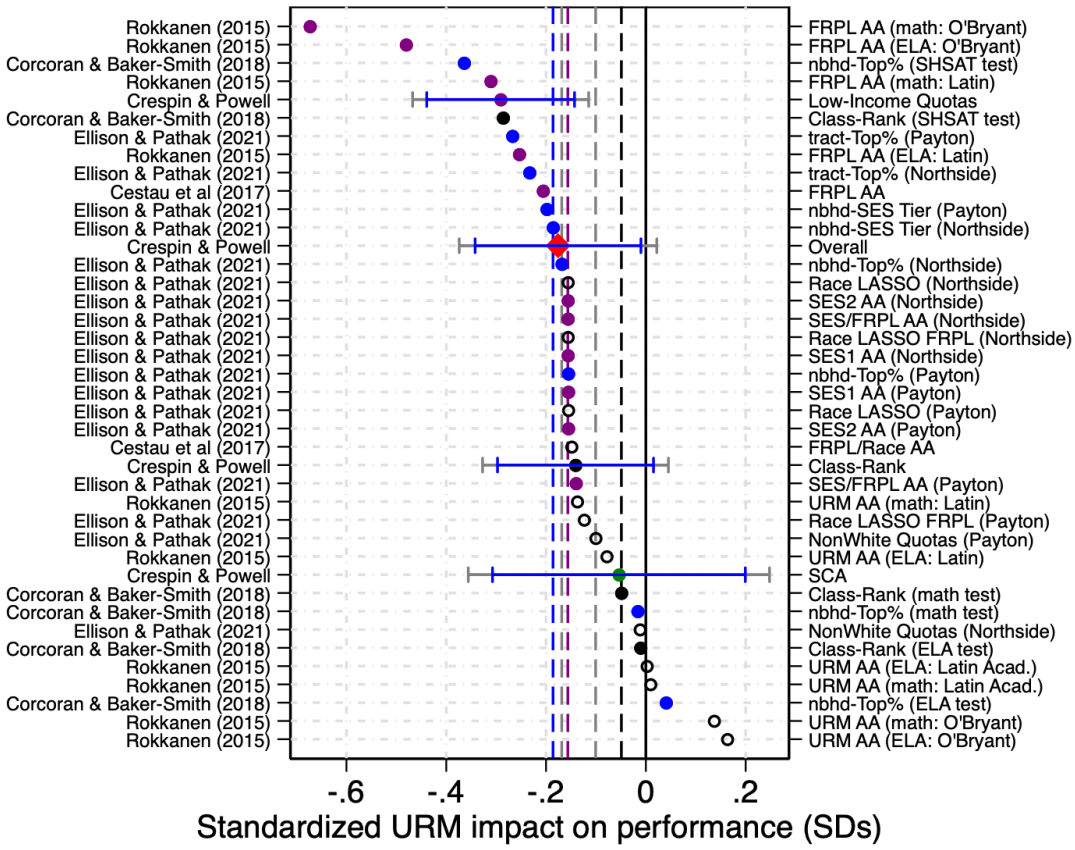
Notes: Based on sample of untreated ASHSs in 2015-16 and treated ASHSs in the year prior to treatment. Each starting point represents each specific ASHS's actual average ACT score for the corresponding group, while the arrow is the simulated average score after matching district URM share.

Figure A22: Forest Plot, FL Enrollment Impacts on Academic Achievement (by specific policy type)



Notes: Each estimate represents the impact of a 10 percentage point increase in FL enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate.

Figure A23: Forest Plot, URM Enrollment Impacts on Academic Achievement (by specific policy type)



Notes: Each estimate represents the impact of a 10 percentage point increase in URM enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate.

Table A1: Summary of K–12 Affirmative Action Simulation Studies

Study	District	Grades	Years	AA Policies	Outcome(s)
Cestau et al. (2017)	Anonymous (midsize urban)	1st	2003–2005	No AA SES-based AA SES- & Race-based AA	SES Composition URM Composition IQ Score
Corcoran and Baker-Smith (2018)	NYC	9th	2009	No AA Middle School Class Rank (10%) Within-Borough Rank (proportional % by borough)	SES Composition URM Composition 7th grade ELA test Avg. 7th grade Math test Avg. ASHS exam Avg.
Ellison and Pathak (2021)	Chicago	9th	2013	No AA Tract SES-Tier Rank Within-Tract Rank (10%) Within-Community Rank (10%) SES-based AA (weighted) SES-based AA (unweighted) SES-based AA (with FRPL) SES-based AA (predicting race) SES-based AA (with FRPL & predicting race) Race Quotas	SES Composition URM Composition ASHS application score
Idoux (2022)	Boston	7th	2004–2016	No AA Middle School Class Rank (5%) Middle School Class Rank (10%)	URM Composition
Rucinski and Goodman (2022)	Boston	7th	2007–2013	No AA Middle School Class Rank (7%) Middle School Class Rank (17%)	SES Composition URM Composition
Rokkanen (2015)	Boston	7th	2000–2004	No AA SES-based AA URM-based AA	SES Composition URM Composition 4th grade ELA test Avg. 4th grade Math test Avg.

Notes: Summarizes the available studies on the effects of race-neutral AA policies on the demographic and academic composition of ASHSs. To be included, a study had to: (a) use data rather than only theoretical models; (b) focus on academically selective K–12 schools; and (c) analyze AA policies that assign admissions priority based on student characteristics (as opposed to modifying application requirements such as GPA or test scores).

Table A2: ASHSs Treated with a Race-Neutral AA Admissions Policy Change in our Analysis Sample

Year	School Name	School District	State	Treated Periods
2019	Brooklyn Latin School	New York City Geographic District #14	NY	4
2019	High School Mathematics, Science, and Engineering at CCNY	New York City Geographic District #5	NY	4
2019	Manhattan/Hunter Science High School	New York City Geographic District #3	NY	4
2019	Queens High School for the Sciences at York College	New York City Geographic District #28	NY	4
2019	Staten Island Technical High School	New York City Geographic District #31	NY	4
2018	Bard High School Early College Queens	New York City Geographic District #24	NY	4
2017	Liberal Arts and Sciences Academy (LASA) High School	Austin Independent School District	TX	4
2016	Academic Magnet High	Charleston County School District	SC	4
2020	Leon M Goldstein High School for the Sciences	New York City Geographic District #22	NY	3
2016	Brooklyn Technical High School	New York City Geographic District #13	NY	3
2016	High School of American Studies at Lehman College	New York City Geographic District #10	NY	3
2021	Beacon High School	New York City Geographic District #3	NY	2
2021	Thomas Jefferson High	Fairfax County Public Schools	VA	2
2021	Boston Latin	Boston Public Schools	MA	2
2021	Boston Latin Academy	Boston Public Schools	MA	2
2021	O'Bryant School Math/Science	Boston Public Schools	MA	2
2019	Lowell High	San Francisco Unified	CA	2
2019	Millennium High School	New York City Geographic District #2	NY	2
2017	Bronx High School of Science	New York City Geographic District #10	NY	2
2022	Townsend Harris High School	New York City Geographic District #25	NY	1
2021	Eleanor Roosevelt High School	New York City Geographic District #2	NY	1
2021	School of the Future High School	New York City Geographic District #2	NY	1
2020	Baruch College Campus High School	New York City Geographic District #2	NY	1
2020	NYC Lab High School for Collaborative Studies	New York City Geographic District #2	NY	1
2018	Stuyvesant High School	New York City Geographic District #2	NY	1
2023	New Explorations into Science, Tech, and Math High School	New York City Geographic District #1	NY	0
2023	Young Women's Leadership School	New York City Geographic District #4	NY	0
2023	Baccalaureate School for Global Education	New York City Geographic District #30	NY	0

Table A3: Summary statistics of ASHSs, subsample with ACT/SAT data (2015–16 SY)

	Full Sample	Final Analysis Sample		
	ASHSs	ASHSs	Control	Treated
	(1)	(2)	(3)	(4)
Free Lunch Shares	0.47	0.38	0.40	0.33
	(0.31)	(0.26)	(0.29)	(0.16)
Asian Shares	0.19	0.22	0.17	0.42
	(0.21)	(0.24)	(0.22)	(0.22)
URM Shares	0.48	0.42	0.49	0.17
	(0.31)	(0.30)	(0.30)	(0.14)
White Shares	0.29	0.32	0.31	0.35
	(0.23)	(0.24)	(0.24)	(0.22)
Other Race Shares	0.04	0.03	0.03	0.05
	(0.03)	(0.03)	(0.02)	(0.04)
9th Grade Enrollment	252.04	256.67	230.51	359.31
	(214.81)	(242.93)	(173.21)	(412.93)
School Enrollment	1056.82	1058.66	966.78	1419.08
	(811.64)	(910.86)	(640.49)	(1571.57)
ACT Score	24.87	25.83	24.47	31.05
	(5.29)	(5.35)	(5.08)	(2.24)
N	117	64	51	13

Note: Summary statistics are based on sample described in Section 2.5. ACT score includes converted SAT scores as described in Section 2.5. Standard deviations are shown in parentheses.

Table A4: Effects of SES-based Admissions Policy on Demographic Enrollment Shares, for balanced sample (initial post-shock period)

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
SES Policy × Post	0.156	0.055***	-0.044	0.006	-0.019
	(0.099)	(0.016)	(0.030)	(0.024)	(0.013)
Percent Effect	46.1%	23.6%	-12.9%	1.7%	-28.6%
Observations	774	774	774	774	774

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A5: Effects on Demographic Enrollment Shares, by policy type

	Low-Income Quotas	Class Rank	Geographic Rank	SCA
	(1)	(2)	(3)	(4)
Panel A: Share Free Lunch				
SES Policy × Post	0.220***	0.079	0.186**	0.074
	(0.055)	(0.055)	(0.089)	(0.055)
Percent Effect	62.1%	50.0%	53.6%	19.6%
Observations	776	714	729	758
Panel B: Share URM				
SES Policy × Post	0.099***	0.068***	0.054	0.041***
	(0.020)	(0.016)	(0.061)	(0.012)
Percent Effect	34.3%	54.4%	16.0%	33.3%
Observations	776	714	729	758
Panel C: Share White				
SES Policy × Post	-0.094**	-0.008	-0.072	0.013
	(0.036)	(0.036)	(0.059)	(0.030)
Percent Effect	-25.8%	-2.0%	-21.4%	4.6%
Observations	776	714	729	758
Panel D: Share Asian				
SES Policy × Post	0.009	-0.040	0.021	-0.013
	(0.035)	(0.050)	(0.015)	(0.047)
Percent Effect	3.4%	-9.7%	7.7%	-2.5%
Observations	776	714	729	758
Panel E: Share Other Race				
SES Policy × Post	-0.014	-0.023*	-0.006	-0.044
	(0.017)	(0.013)	(0.023)	(0.033)
Percent Effect	-19.2%	-41.8%	-11.1%	-64.7%
Observations	776	714	729	758

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A6: Effects of SES-based Admissions Policy on Enrollment Shares, excluding or including NYC schools

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
Panel A: Treated units only include non-NYC academically selective high schools					
SES Policy Change X Post	0.158***	0.050	-0.020	-0.013	-0.020
	(0.056)	(0.036)	(0.019)	(0.026)	(0.015)
Percent Effect	79.1%	18.9%	-5.5%	-4.2%	-36.0%
Observations	748	748	748	748	748
Panel B: Treated units only include NYC academically selective high schools					
SES Policy Change X Post	0.105**	0.068***	-0.045	0.001	-0.025
	(0.051)	(0.016)	(0.033)	(0.030)	(0.021)
Percent Effect	27.3%	30.7%	-13.5%	0.3%	-35.6%
Observations	843	843	843	843	843

Notes: Estimation is based on equation 1. Panel A excludes treated New York City ASHSs, while Panel B excludes treated non-New York City ASHSs. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A7: Robustness to Sample Choices

	Main Estimates	Pre-2020	Pre-2021	No 100% FL Years	No 100% FL Schools	Only Top ASHSs	No Treated District Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share FL	0.135**	0.088	0.124**	0.104*	0.078	0.163**	0.147**
	(0.056)	(0.060)	(0.060)	(0.056)	(0.053)	(0.066)	(0.059)
Share URM	0.065***	0.057***	0.061***	0.064***	0.061***	0.070**	0.066***
	(0.017)	(0.013)	(0.015)	(0.017)	(0.017)	(0.035)	(0.018)
Share White	-0.042	-0.028	-0.047**	-0.043	-0.038	-0.028	-0.040
	(0.026)	(0.022)	(0.023)	(0.026)	(0.026)	(0.023)	(0.026)
Share Asian	-0.001	-0.005	0.006	0.001	-0.000	-0.017	-0.002
	(0.022)	(0.017)	(0.020)	(0.022)	(0.022)	(0.032)	(0.022)
Share Other	-0.024	-0.027	-0.024	-0.024	-0.026	-0.029*	-0.025
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.017)	(0.019)
N	898	620	703	878	827	749	838

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. The sample in column (2) excludes all cohorts after 2019. The sample in column (3) excludes all cohorts after 2020. The sample in column (4) drops any observations with 100 percent share of students eligible for Free Lunch. The sample in column (5) drops any school that has at least one year with 100 percent share of students eligible for Free Lunch. The sample in column (6) keeps only the highest-ranking ASHS within each district when multiple ASHSs are treated within the district (effectively drops non-Stuyvesant ASHSs in NYC and non-Boston Latin ASHSs in Boston). The sample in column (7) drops all control ASHSs in districts with treated ASHSs from the analysis sample. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A8: Robustness to Additional Controls

	Main Estimates	Race Controls	Income Controls
	(1)	(2)	(3)
Share FL	0.135**	0.149***	0.133**
	(0.056)	(0.054)	(0.056)
Share URM	0.065***	0.063***	0.064***
	(0.017)	(0.017)	(0.017)
Share White	-0.042	-0.045*	-0.042
	(0.026)	(0.025)	(0.026)
Share Asian	-0.001	-0.001	-0.001
	(0.022)	(0.022)	(0.022)
Share Other	-0.024	-0.018	-0.024
	(0.018)	(0.015)	(0.018)
N	898	875	890

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. Column (2) includes controls for yearly changes in 9th-grade racial shares for the entire district. Column (3) includes controls for annual changes in 9th-grade low-income shares for the entire district. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table A9: Robustness to Alternative Specifications

	Main Estimates	Stacked DiD	No De-Trending
	(1)	(2)	(3)
Share FL	0.135**	0.081***	0.045
	(0.056)	(0.023)	(0.029)
Share URM	0.065***	0.054***	0.028**
	(0.017)	(0.015)	(0.012)
Share White	-0.042	-0.039	-0.025
	(0.026)	(0.024)	(0.019)
Share Asian	-0.001	-0.003	-0.005
	(0.022)	(0.022)	(0.015)
Share Other	-0.024	-0.010	0.002
	(0.018)	(0.013)	(0.004)
N	898	3615	898

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. The model used for column (2) is a stacked difference-in-difference approach that creates a sub-experimental dataset with centered control observations for each treatment year, then estimates a weighted regression with the main model (Wing et al., 2024). The model for column (3) does not include school-level pre-period de-trending. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A10: Effects on Academic Performance, by policy type

	Avg. ACT Score		
	LIQ policies	Class Rank policies	SCA policies
	(1)	(2)	(3)
Admissions Change	-1.425***	-0.360	-0.110
	(0.482)	(0.238)	(0.313)
Percent Effect	-5.0%	-1.2%	-0.3%
N	439	421	465

Notes: Estimation in panel A is based on equation 1. Panels B and C report 2SLS coefficients where SES Policy implementation is the instrument and either FL Enrollment Shares or URM Enrollment Shares is the endogenous variable, respectively. IV estimates are multiplied by 10 for ease of presentation. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in section 2. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A11: ACT Impacts - Robustness to Sample Choice and Specification

	Main Estimate	Non-NYC Treatments Only	NYC Treatments Only	Stacked DiD
	(1)	(2)	(3)	(4)
SES Policy Change X Post	-0.529	-0.913*	-0.415	-0.999***
	(0.326)	(0.505)	(0.353)	(0.243)
Percent Effect	-1.8%	-3.2%	-1.4%	-3.5%
Observations	501	429	484	1449

Notes: Estimation is based on equation 1, using an unbalanced sample. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in Section 2. Column (1) replicates the main ACT estimates. In column (2) the treated units only include NYC ASHSs. In column (3) the treated units only include non-NYC ASHSs. Estimates in Column (4) are from a stacked difference-in-difference approach that creates a sub-experimental dataset with centered control observations for each treatment year, then estimates a weighted regression based on Wing et al., 2024. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A12: IV Estimates of Effects on Academic Performance

	Avg. ACT Score
	(1)
Panel A: IV estimates (SES composition)	
FL Enrollment Shares	-0.959
	(0.593)
Percent Effect	-3.1%
Standardized Effect	0.18σ
Observations	501
Panel B: IV estimates (racial composition)	
URM Enrollment Shares	-0.940*
	(0.538)
Percent Effect	-3.0%
Standardized Effect	0.18σ
Observations	501

Notes: Estimation in panel A is based on equation 1. Panels B and C report 2SLS coefficients where SES Policy implementation is the instrument and either FL Enrollment Shares or URM Enrollment Shares is the endogenous variable, respectively. IV estimates are multiplied by 10 for ease of presentation. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in section 2. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Appendix B Policy Changes

B.1 Low-Income Student Quota Admissions Policy

B.1.1 New York City Diversity in Admissions Program (New York, NY)

New York City has sought to increase diversity in its non-specialized schools through prioritization of students with free or reduced-price lunch status. In 2012, PS 133 William A. Butler School in Brooklyn became the first to use this approach, setting aside 35 percent of its seats for FRPL and English language learner (ELL) students.[56]

In the years following the implementation of this policy at PS 133, pressure mounted on the city to make efforts to integrate its public school system. A report from the UCLA Civil Rights Project found that the state of New York had the most segregated school system in the United States.[57] In 2014, a coalition of students founded IntegrateNYC—a non-profit which has since filed a lawsuit against New York City's admissions process.[58] Additionally, the state education department enacted the Socioeconomic Integration Pilot Program (SIPP) which offered up to $10 million from 2015-2018 to support progress toward integration plans and implementation at Title I focus schools in the state.[59]

Beginning with the 2016-17 school year, New York City launched the Diversity in Admissions pilot program at seven schools. Under this program, these schools set aside seats for priority groups such as those eligible for free or reduced-price lunch and English language learners.[60] District leaders were not allowed to explicitly target race in admissions because of the 2007 Parents Involved SCOTUS case. The NYC Education Department followed up the Diversity in Admissions pilot with a diversity plan[61] in June 2017 that set out the city's first publicized diversity goals.[62] One tenet of the city's diversity plan was to expand the Diversity in Admissions program. It proposed outreach to school leadership and communities to spread awareness of the program and its application process.

Building on the pilot program of 2016-17, the Diversity in Admissions program grew from seven to 19 schools for the 2017-18 school year. An additional 11 schools joined the

[56] https://www.dnainfo.com/new-york/20141210/park-slope/how-park-slope-school-is-fighting-stay-diverse/

[57] https://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/ny-norflet-report-placeholder/Kucsera-New-York-Extreme-Segregation-2014.pdf

[58] https://ny.chalkbeat.org/2021/3/9/22321193/nyc-students-file-lawsuit-taking-aim-at-admissions-screens-school-segrega

[59] https://www.p12.nysed.gov/funding/2015-18-title-1-ses-integration-grant/ses-integration-grant.pdf

[60] https://static1.squarespace.com/static/53ee4f0be4b015b9c3690d84/t/5c4f765088251bb7f8e4fd6c/1548711510681/PromisingOutcomes.pdf

[61] https://web.archive.org/web/20170611214002/http://schools.nyc.gov/NR/rdonlyres/D0799D8E-D4CD-45EF-A0D5-8F1DB246C2BA/0/diversity_final.pdf

[62] https://www.nytimes.com/2017/06/06/nyregion/long-awaited-plan-for-integrating-schools-proves-mostly-small-bore.html

program in 2018-19, including the first academically-selective high school in NYC to set aside seats for FRPL-eligible students—Bard High School Early College Queens. The program has continued to grow, with 43 high schools participating in 2023-24.[63] This includes 13 schools that are documented as ASHSs by Finn and Hockett (2012).

Mader et al. (2018) provide a descriptive analysis of the changes in enrollment demographics in participating schools in the first two years of the Diversity in Admissions Initiative (none of the schools during this time were ASHSs). They showed significant increases in low-income enrollment, but no statistically significant change in any race/ethnicity category on average across the pilot schools.

B.1.2 Liberal Arts and Science Academy (LASA) High School (Austin, TX) – Diversity Process

Until 2017, LASA, located in Austin, Texas, had admitted students based solely on their performance on the school's admissions exam, past grades, an application essay, and teacher recommendations. For the cohort applying for fall 2017 admission, the school considered applicants' race and socioeconomic status for the first time. At this point, LASA shifted to an admissions model in which 80 percent of students were admitted based on their holistic application score. The remaining 20 percent of admitted students are chosen based on their application scores, contingent on meeting at least three out of the five special circumstances defined by the school. These circumstances are being a future first-generation college student, being zoned to a neighborhood middle school with at least 70 percent economically disadvantaged students, belonging to an underrepresented racial or ethnic group, demonstrating academic achievement in the face of difficult circumstances, and embracing academic challenges.[64]

In addition to amending the admissions criteria, LASA also sent out a recruitment team made up of parents of LASA students during the 2017 admissions season. This group met with students at all Title I middle schools in the district and provided guidance to those seeking admission to LASA.[65]

B.2 Middle School Class Rank Priority

Another type of policy mechanically increases the geographic diversity of admitted students by automatically granting admission to a specified number of students from each middle

[63] https://www.schools.nyc.gov/enrollment/enrollment-help/meeting-student-needs/diversity-in-admissions

[64] https://lasa.austinschools.org/acceptance-procedures

[65] https://www.kxan.com/news/lasa-aims-to-diversify-campus-with-new-recruiting-application-process/

school in a district. This approach dictates that students from every school have an opportunity to be admitted to the selective high school.

B.2.1 Academic Magnet High School (North Charleston, SC) – Top Two Program

Academic Magnet High School in North Charleston, South Carolina implemented this approach through a "Top Two Program" that first applied to the fall 2016 cohort.[66] In this admission system, Academic Magnet High accepts the top two students, by grade point average, from every public middle school in the district. The offer of admission is conditional on the students achieving a minimum score on the school's application rubric which includes grades, standardized test scores, a math task, a writing task, and a teacher recommendation.

B.2.2 Thomas Jefferson High School (Alexandria, VA) – Top 1.5% Policy

In December 2020, the Fairfax County School Board voted to change the admissions policy for Thomas Jefferson High School for Science and Technology. At this meeting, two alternative methods for selecting students were proposed. First, a proposal to enact a Hybrid Merit Lottery—which would have admitted the top 100 scoring applications, while allocating the remaining 450 seats by lottery among qualified applicants—was rejected. The Board then approved an admissions policy that admits the top 1.5 percent of students from each middle school based on their application scores. This leads to an allocation of approximately 450 seats, with the remaining 100 seats filled by the highest-scoring applicants who have yet to be admitted. This policy has been subject to several recent legal challenges, with the U.S. Court of Appeals for the Fourth Circuit finding in favor of the School Board in May 2023.[67] There have been calls for the case to be heard by the Supreme Court.[68]

B.3 Geographic Rank Priority

B.3.1 Lowell High School (San Francisco, CA) – Priority Middle School

Lowell High School began its current practice of setting aside 15 percent of seats for students from underrepresented middle schools in 2001. In October 2018, the San Francisco Unified School District Board of Education approved an update to the eligibility criteria under this third band of their admissions system.

[66] https://www.postandcourier.com/news/academic-magnet-continues-offering-admissions-to-top-2-graduates-of-middle-s article_65c57dbe-e250-11e6-9a39-171418065599.html

[67] https://www.ca4.uscourts.gov/opinions/221280.P.pdf

[68] https://www.nytimes.com/2023/08/21/us/politics/supreme-court-thomas-jefferson-high-school-admissions.html

Lowell analyzes the enrollment totals at each middle school in San Francisco, total enrollment in the district, and the number of applicants and admitted students from each school over the three preceding years. With these numbers, Lowell identifies a subset of middle schools that are underrepresented within Lowell High School. Under this framework, if an underrepresented school becomes especially proficient at placing students at Lowell, they can lose their underrepresented status. This update identified Willie Brown Middle School as a Band 3 school, regardless of its students' representation within Lowell. The change went into effect for the cohort that entered Lowell for the 2019-20 school year.[69]

It was estimated that the policy would affect 11 current 8th graders at Willie Brown when it was enacted in October 2018.[70] The policy could also impact the composition of Willie Brown as middle school parents respond to the policy. For example, from 2020-2022, the share of Black students at Willie Brown has fallen from 33 percent to 21.9 percent and the share of white students has risen from 10.5 to 18.0 percent. The share of students qualifying for Free or Reduced-price Lunch fell from 66.5 to 52.0 percent.[71]

Lowell instituted lottery admissions which affected the 2021-22 and 2022-23 entering cohorts before returning to selective admissions for 2023-24. Therefore, we only estimate the effects of the 2019 policy change on the 2019 and 2020 cohorts.

B.3.2 Goldstein High School for the Sciences (New York, NY) – Priority for Underrepresented Zip Codes

In 2019, Leon M. Goldstein High School for the Sciences planned to start a Diversity in Admissions Initiative, which would be implemented starting in 2020.[72] Their initiative was to prioritize students from underrepresented zip codes. To do this, they rank order students within each zip code and admit the top ranking students in each zip code, improving the odds of admissions for students who were historically less likely to be admitted.[73]

B.3.3 Boston Public Schools – Socioeconomic Tiers Based on Address

Boston Public Schools operates three academically selective high schools: Boston Latin, Boston Latin Academy, and the O'Bryant School of Mathematics and Science. In 1997, these schools were legally mandated to remove their racial quotas for incoming cohorts. Starting with the fall 2021 cohort, BPS instituted a new admissions system that divides the city's

[69] https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=AZS28M7BDEA9

[70] https://www.sfchronicle.com/education/article/SF-school-board-changes-Lowell-admission-policy-13294817.php

[71] https://drive.google.com/drive/folders/1MQg9Lrdw0xKxc8FZ8eD2No67kB5P4rYq

[72] https://www.nydailynews.com/2019/10/02/four-nyc-high-schools-adopt-new-admissions-methods-to-increase-diversity

[73] https://www.myschools.nyc/en/schools/high-school/22K535/

Census tracts into eight socioeconomic tiers. Applicants continue to be assessed based on their middle school grades and test scores; however, they are only compared to other students within their tier. An equal number of seats are allocated to each SES tier.[74,75]. Additionally, extra application score points are assigned to students experiencing homelessness, foster care, or public housing, as well as students that in the prior year attended a school that serves at least 40% low-income students. For fall 2021 only, the entrance exam was not used due to the COVID-19 pandemic, but the test was reinstituted, along with the socioeconomic tier system beginning with the fall 2022 cohort.[76] For the 2024 cohort, the extra points for students who attended a school with a high proportion of low-income students will depend on the tier the student lives in.[77]

B.4 Submarginal Conditional Acceptance (SCA)

B.4.1 New York City Specialized High School Discovery Program

The NYC Specialized High School Discovery Program is an alternative admissions option for disadvantaged students who scored within a certain range below the admissions cutoff on the Specialized High School Admissions Test (SHSAT). The main goal of this program is to increase diversity in New York's specialized high schools. Historically, specialized school participation in this program has been optional. A provision in the 1971 New York State law that required admission to the city's specialized high schools to be based solely on the admissions test allowed specialized high schools to sponsor a Discovery Program [78]. By 2015, the majority (5/8) of specialized high schools were still admitting some students through the Discovery Program. The exceptions were Stuyvesant, Bronx Science, and the High School for American Studies at Lehman College.[79,80] At that time, students were eligible for the program as long as they scored within a certain range below the admissions cutoff, were recommended by their middle school, attended and passed a summer preparatory program administered by the specialized high school, and were certified as disadvantaged, based on household income, family receipt of government aid, foster care status, or being recent immigrants who are classified as English language learners.

[74] https://www.bostonpublicschools.org/tier-map

[75] https://www.wbur.org/news/2021/07/15/new-exam-school-admissions-policy

[76] https://www.the74million.org/article/scotus-passes-on-hearing-affirmative-action-case-involving-elite-boston-schools/

[77] https://web.archive.org/web/20241008025235/https://drive.google.com/file/d/1i8R0AJO2o1wtlt7TrTRVMZX8UTyK24xs/view

[78] https://web.archive.org/web/20140623060948/http://nypress.com/benign-neglect/

[79] https://www.nytimes.com/2018/08/13/nyregion/discovery-program-specialized-schools-nyc.html

[80] Since the early 2000s, Stuyvesant and Bronx Science had terminated their participation in the Discovery Program https://web.archive.org/web/20140623060948/http://nypress.com/benign-neglect/.

During the 2016-2017 school year, the NYC DOE released a diversity plan that included their goals and existing strategies to increase diversity in NYC schools. Part of this plan was to expand the Discovery plan. This included restarting the Discovery program at the High School for American Studies at Lehman College in the summer 2016, at Bronx Science in summer 2017, and at Stuyvesant in summer 2018. Additionally, in summer 2016, they expanded the number of Discovery Program students admitted at Brooklyn Tech. Beginning in summer 2019, Discovery Program eligibility was restricted to require that all students have attended a high-poverty middle school in addition to being disadvantaged (low-income family, ELL student recently moved to NYC, and/or student in temporary housing). By 2019-20, 20 percent of seats at each specialized high school were reserved for Discovery Program participants.[81]

[81] https://www.schools.nyc.gov/docs/default-source/default-document-library/specialized-high-schools-proposal

Appendix C Addressing Multiple Policy Changes

To identify the effects of specific SES-based admissions policies, we must ensure that we are capturing "clean" policy changes. This mean that if many policy changes are happening at once, it would be difficult to identify which of those policy changes led to the outcome of interest. Therefore, we must deal with issues of this type when we're creating our analytic dataset of policy changes.

Nine ASHSs in Chicago were removed due to transitioning from a race-based affirmative action admissions policy to a place-based admissions policy. While we would be interested in the impacts of the SES place-based policy in Chicago, we believe the termination of the race-based admissions policy contemporaneously impacts admissions in a way that obscures the impact of the SES-based policy.

Additionally, before compiling our data to create Table 1, we removed schools with policy changes in consecutive school years that are impossible to disentangle from each other, as each new policy change may have a new intention/goal. This decision leads us to drop: Academy at Palumbo, Carver High School, Central High School, Julia R Masterman Secondary School, Parkway Center City Middle College, Paul Laurence Dunbar High School, Western High School, and Maggie L. Walker Governor's School.

In some cases, consecutive policy changes have the same intention/goal. We merge policy changes that occur in consecutive school years to be represented as a single policy change in the first year of incidence because all the follow-up changes were made to strengthen the adjusted policy using the same strategy as the initial change. These schools are: McKinley Technology High School, School Without Walls High School, Science Park High School, Baruch College Campus High School, Beacon High School, NYC Lab High School for Collaborative Studies, and School of the Future High School.

As described in section 2, we further restrict our analytic sample of treated ASHSs to those that are not contaminated by multiple policy changes. We allow ASHSs that had multiple admissions policy changes, as long as they meet our requirement that the treatment timing does not overlap with a previous or future admissions policy change at the same school. That is, we require that the policy change must be at least six years after any previous policy change and at least one year prior to any future policy change. In practice, no school appears more than once in our analysis sample, though many schools have multiple policy changes.

Appendix D Dose-Response vs. Traditional DID and Event Study Estimates

As described in section 2, our socioeconomic enrollment data is only available at the school-by-year-level. Our main empirical approach incorporates dose-response to address the fact that our socioeconomic unit of observation is only partially impacted in the first three years that the policy is implemented. With proper scaling, this approach is mathematically equivalent to the traditional approach without dosage. Furthermore, our preferred approach provides estimates that are more easily interpretable and comparable than the traditional DID approach, given the limitations of the socioeconomic enrollment data.

Appendix Figure D1 shows an event study that is analogous to that of figure 4, but without dose-response. The event study shows a small, but immediate, deviation at period zero that increases over time. However, each post-shock coefficient has a different interpretation since only a portion of the school is treated and, hence, must be rescaled for comparability.

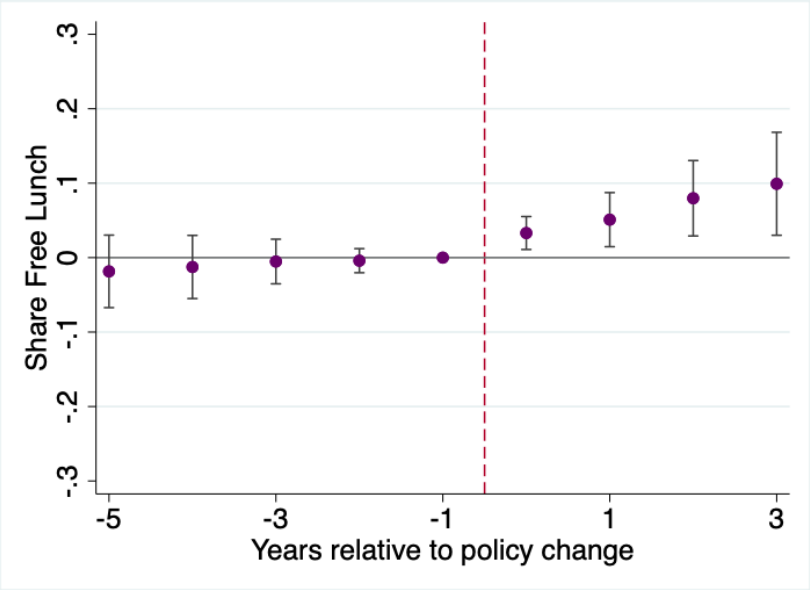
Appendix Table D1 shows that DID estimates that do not incorporate a dose-response are mathematically equivalent to those calculated by the dose-response DID approach as long as they are rescaled appropriately. Column (1) shows dynamic DID coefficients estimated in each post-shock period without accounting for dosage. Column (2) shows the proportion of the typical high school that is treated with the additional cohort entering under the new admissions policy. In column (3), we rescale the corresponding estimate by the proportion of cohorts treated, which provides the estimated impact if the whole unit of observation had been treated. In column (4), we show the estimated dynamic impacts from our main approach that incorporates the dose-response. Column (1) and (4) show that both approaches provide estimates with the same level of statistical significance.

Table D1: Equivalence between dynamic DID estimates, without and with dose-response

	Estimates w/o Dose-Response	Proportion Treated	Rescaled Estimates	Estimates w/ Dose-Response
	(1)	(2)	(3)	(4)
Panel A: Impact on Free Lunch Enrollment Share				
SES Policy Change X Post0	0.028*	1/4	0.112	0.111*
	(0.014)			(0.057)
SES Policy Change X Post1	0.048**	2/4	0.096	0.095**
	(0.021)			(0.042)
SES Policy Change X Post2	0.077***	3/4	0.103	0.103***
	(0.027)			(0.036)
SES Policy Change X Post3	0.102***	4/4	0.102	0.102***
	(0.038)			(0.038)
Observations	898			898

Notes: Estimates are based on dynamic DID specification, which allows for a separate estimate in each post-shock period. We keep observations with relative time up to six periods after treatment (or after 2016 for control units) in the analysis sample, and we bin relative time periods greater than or equal to four. Refer to section D for details. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Figure D1: Event Study Estimates (without dose-response) – Free-Lunch Enrollment Shares



Notes: Event study is based on equation 2, but without dose-response, therefore $AdmissionsChange_{st}$ is always set to 1 for all treated units. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend.

Appendix E Secular Trends

As discussed in section 4.4, a potential concern with our identification strategy is the existence of secular trends that differ systematically with treatment exposure. The event study estimates suggest such trends are not biasing our estimates. As an additional test of whether the timing pattern of the AA admissions policy change is driving our results we perform permutation tests.

We conduct placebo tests where we randomly assign treated ASHSs to implementation years and test whether the new estimated effects are as large as our baseline policy impact estimates. In Table E1, we distribute the treated ASHSs equally across our implementation years between 2016 and 2022.

We randomize placebo treatment dates 100 times for each outcome and calculate the percent of cases that generate policy impact estimates larger than our baseline estimates. Therefore, these results represent p-values for the null hypothesis that any combination of policy implementation dates would generate as extreme of policy impact estimates as our results. In Table E1, we reject the null of such a hypothesis for estimates of Free Lunch share, URM share, and White enrollment share. We do not reject the null of such a hypothesis for Asian or other race enrollment shares. This is consistent with our main estimates for Asian and other race enrollment shares because they are not statistically significant.

Table E1: p-values from Permutation Tests—100 iterations

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
Percent greater than baseline	0.06	0.04	0.05	0.73	0.12

Notes: We randomize placebo treatment dates 100 times for each outcome and calculate the percent of cases that generated policy impact estimates larger than our baseline estimates.

Appendix F Algebra 1 Regents Exam Data

We measure incoming (or early) student achievement for a subset of schools using school-by-year New York State Regents exam results from 2012–13 through 2023–24.[82] These data report, for each school and year, the number of test-takers for each subject exam in each demographic group (overall, by SES, and by race) who fall into each performance category for each subject. We focus on Algebra I, the lowest-level math Regents exam that is offered for high schoolers.

New York State revised its Regents exams and performance categories during our study period. Prior to 2013–14, scores were grouped into four categories: 0–54, 55–64, 65–84, and 85–100. Beginning in 2013–14, the state introduced Common Core–aligned Regents exams—still scored 0–100 but divided into five achievement levels: Level 1 (0–54), Level 2 (55–64), Level 3 (65–74), Level 4 (75–84), and Level 5 (85–100). Math exams transitioned between 2013–14 and 2015–16, during which students could take either the Common Core or prior version; by 2016–17 only the Common Core version remained.[83]

To create a consistent measure over time, we map all exams into the original four-level scale, combining Common Core Levels 3 and 4 into a single category. Our analysis focuses on the top performance category, whose score range (85–100) is stable across all years. When both Common Core and non–Common Core Algebra I exams were offered in the same year (e.g., 2013–14), we aggregate the number of students in each performance category across the two exams. Finally, because the data include the number of students taking each Regents exam, we treat these counts as proxies for the number of students enrolled in the corresponding course in that year.

[82] https://data.nysed.gov/downloads.php

[83] https://www.regents.nysed.gov/sites/regents/files/meetings/CommonCoreGeoExams.pdf